Bonn-Aachen International Center for Information Technology (B-IT)
University of Bonn
Master Program in Life Science Informatics

MASTER'S THESIS FOR OBTAINING THE ACADEMIC DEGREE
„MASTER OF SCIENCE (M.SC.)“

Extracting Most Predictive Subgraphs From Models of Human Brain Connectivity

Author:	*First Examiner:*
Shreya Kapoor	Prof. Dr. Thomas Schultz
	Second Examiner:
	Prof. Dr. Holger Fröhlich
	Advisor:
	Mohammad Khatami

Submitted: November 16, 2020

Abstract

Advances in non-invasive neuroimaging modalities have led to a growing interest in investigating the underpinnings of human cognition and neurological disorders. Analysis of human brain connectivity or the human connectome has hence become an active area of research. Computational methods based on graph theory have played a significant role in understanding the topological organization of brain networks. Simultaneously, machine learning algorithms are now being used in neuroimaging studies to classify diseased states in patients. In this work, we propose a method to combine the utilities of graph theoretic methods with machine learning classifiers to obtain discriminative subnetworks of the brain. The novel component of this thesis is to reduce the graphical network derived from structural brain connectivity. Subnetworks were extracted by approximating the solution to the Maximum Edge Weight k-Induced Subgraph (MEWIS) problem that falls into the class of NP-hard problems. The use of MEWIS as a feature selection technique makes our classification results more interpretable as compared to traditional filter-based methods. The experimental results from our analysis show that MEWS subgraphs can achieve 95% area under ROC curve (AUC) for gender classification. Furthermore, it is a generalized method that is independent of the classification task and can hence be useful for the discovery of disease biomarkers.

Contents

List of Figures

List of Tables

list

Acronyms

5TT	Five-Tissue type
AC-PC	Anterior and Posterior Commissure
ACT	Anatomically Constrained Tractography
ADC	Apparent Diffusion Coefficient
ANN	Artificial Neural Network
API	Application Processing Interface
AUC	Area Under ROC Curve
CSD	Constrained Spherical Deconvolution
CSF	Cerebro-Spinal Fluid
CT	Computed Tomography
dMRI	Diffusion MRI
DTI	Diffusion Tensor Imaging
DWI	Diffusion Weighted Imaging
FA	Fractional Anisotropy
fMRI	Functional Magnetic Resonance Imaging
fODF	fiber Orientation Distribution Function
FOV	Field of View
GM	Gray Matter
GMWCS	Generalized MWCS
HARDI	High Angular Resolution Imaging
HCP	Human Connectome Project
MD	Mean Diffusivity
MEWIS	Maximum Edge Weight k-Induced Subgraph
MIP	Mixed Integer Programming

MLP	Multi Layer Perceptron
MRI	Magnetic Resonance Imaging
MSMT-CSD	Multi-Shell Multi-Tissue Constrained Spherical Deconvolution
MWCS	Maximum Edge Weight Subgraph
NMR	Nuclear Magnetic Resonance
PCA	Principal Component Analysis
PDF	Probability Distribution Function
PET	Positron Emission Tomography
PGSE	Pulsed-gradient Spin-echo
RF	Radio Frequency
RF	Random Forests
RGB	Red Green and Blue
ROI	Region of Interest
SD	Spherical Deconvolution
SE	Single-shot Spin-echo
SIFT	Spherical-deconvolution Informed Filtering of Tractograms
SNR	Signal-to-Noise Ratio
SVC	Support Vector Classifier
SVM	Support Vector Machines
TE	Echo Time
TR	Repetition Time
WM	White Matter

1 Introduction

One of the greatest scientific challenges in the 21^{st} century is to understand the structure and function of the human brain. The mechanism in which intricate networks of cells organize themselves at different scales and give rise to behavior, emotion, and intelligence has remained a topic of discussion in neuroscience for ages. In recent years, there has been a surge of interest to understand the human brain in health and disease. Technological advances are taking humanity closer to answering the question: what makes us, us?

Research in Neuroscience has greatly benefited from advances in neuroimaging data acquisition, analysis, and retrieval. The decreasing cost of data storage and computational memory resources has lead to the development of imaging standards and made a wide range of studies possible. Increasing efficiency and accuracy of non-invasive imaging modalities such as MRI, Computed Tomography (CT), and Positron Emission Tomography (PET) have contributed significantly towards understanding the human brain in health and disease. Among these, MRI is a popular modality due to the diversity of information that can be generated using a single scanner. Each it's four types: structural MRI, task-activated functional MRI, diffusion MRI, and resting-state functional Magnetic Resonance Imaging (fMRI) can be used to carry out different types of research (Van Essen and Glasser, 2016).

The excitement from investigators has brought about the progress of ambitious projects such as the HCP launched in 2009 (Van Essen and Glasser, 2016). The HCP explores the research area known as *connectomics* (previously known as *hodology*), the study of the brain's structural and functional networks. Connectomics aims at creating a map of the brain that is divided into functionally and structurally distinct areas known as parcels. Understanding the contribution of connections between regions is crucial to analyze the emergent properties of brain structure and function. This framework is also important to understand defects caused due to neurological disorders as well as emotion and cognition (Sala-Llonch et al., 2015).

Structural brain connectivity is the study of white matter tracts in the brain and can be measured using Diffusion Tensor Imaging (DTI). Functional connectivity is based on information exchange between different brain regions and can be measured using fMRI. Each of these connectivities has its importance. However, DTI based structural connectivity is of particular importance for medical applications as it provides insight into anatomical connections within an individual's brain, and can aid precision medicine (Cociu et al., 2017).

Machine Learning has become an indispensable tool in neuroimaging due to the *curse of dimensionality* in such studies. Here, the number of subjects in the study is often much lesser than the number of features. It is often used to make predictions based on MRI scans and to classify diseased states of patients. However, for the analysis of brain connectivity there is a need for the machine learning algorithm to understand the topological properties of brain organization. Generic classifiers such as Naive Bayes are not well suited for brain networks due to the density and computational complexity of such networks. Traditional classifiers suffer from a lack of interpretability and often do not lead to results with neurobiological correspondence (Sarica et al., 2017). Hence, it is required to incorporate the advantages of graph theoretic methods and machine learning algorithms for making neurobiologically and statistically significant predictions.

2 Background

Non-invasive neuroimaging techniques have become indispensable tools to understand cognitive processes and neurological disease pathology. Diffusion MRI (dMRI) is a new type of imaging modality often used to produce highly detailed anatomical images. It is suitable for looking at soft tissues such as those in the nervous system. White matter tractography from dMRI scans enables the inference of structural connectivity patterns between brain regions. Analysis of individual differences in structural brain connectivity is relevant to understand emotion, behavior, and cognition.

Individual differences in neuroimaging are often evaluated using machine learning classifiers. Classification based on neuroimaging techniques is an intricate task. It suffers from the *curse of dimensionality* and lack of interpretability. Using graph based techniques along with feature selection methods can help establish neurobiological correspondence of the decision made by a classifier.

This chapter describes the prerequisites required to understand the implementation in chapter 3. In the first few sections, the technical details of imaging using MRI and dMRI will be presented. Subsequent sections give an introduction to the prior work relevant for classification based on structural brain connectivity.

2.1 Magnetic Resonance Imaging

MRI is a non-invasive imaging technology that can produce 3-dimensional detailed anatomical images (McRobbie et al., 2006). This technique is based on the Nuclear Magnetic Resonance (NMR) principle. NMR a physical phenomenon in which nuclei respond to a combination of a constant and a weakly oscillating magnetic field by producing a signal at their resonant frequency. The difference between the magnetic properties of different tissue types subjected to an external magnetic field is used to generate images of biological specimens. Such electromagnetic properties are used to form images since the human body is 65% water (H_2O). Water has an innate dipole moment of $1.84D$ and its hydrogen nuclei act like little magnets since they have an odd number of protons and non-zero net spin.

In normal conditions (Fig. 2.1.a), i.e. in the absence of an external magnetic field the spins of the hydrogen nuclei are randomly oriented and appear to cancel out each other when observed as an ensemble. Once the external field is applied, the hydrogen nuclei exhibit their paramagnetic nature and gain a net magnetization in the direction of the external magnetic field. This magnetization is called as the longitudinal magnetization. The initial longitudinal magnetization can be represented using the following equation:

$$\vec{M_o} = \mu \vec{B_0} \tag{2.1}$$

where M_0 is the net magnetization, μ is the magnetic permeability or the inertia of a material to get magnetized and B_0 is the external magnetic field. According to Fig. 2.1.a it can be seen that the direction of B_0 defines the coordinate system for the imaging experiment. The transverse plane is defined as the plane perpendicular to the direction of B_0. The magnetization in this plane is termed as the transverse magnetization. In the presence of the external static magnetic field, the individual nuclei are precessing (rotating around their axis) with a frequency known as the Larmour frequency, which can be expressed by the equation:

$$\omega_0 = \gamma B_0 \tag{2.2}$$

where ω_o represents the frequency, γ represents the gyromagnetic ration and B_0 the external field.

An oscillating field operating at the Larmour frequency Eq. (2.2) is then added to the static magnetic field using an radio frequency (RF) coil. This field is termed as B_1 in Fig. 2.1.a. The B_1 field causes the direction of the net magnetization vector to get altered i.e. ‘flipped’ or ‘tipped’ out of alignment with B_0. The net magnetization gets tipped towards B_1 with the angle of rotation termed as the flip angle. The RF pulse exerts a torque which can be mathematically expressed as:

$$\vec{\tau} = \vec{m} \times \vec{B_1} \tag{2.3}$$

where $\vec{m}$ is the magnetic moment and the $\vec{B_1}$ is the applied magnetic field. The flipping does not bring all the spins in phase with each other but causes the net magnetization to flip onto the transverse plane (M_{xy} in Fig. 2.1.a). The magnetization M_{xy} then precesses in this plane. It is important to note that the net magnetization does not precess until an external force disturbs its equilibrium position of alignment with the static magnetic field. When the magnetization precesses, magnetic resonance is said to occur.

After this, the RF pulse is switched off which causes the tissue molecules to return to their original state and is termed as the relaxation phase. The relaxation

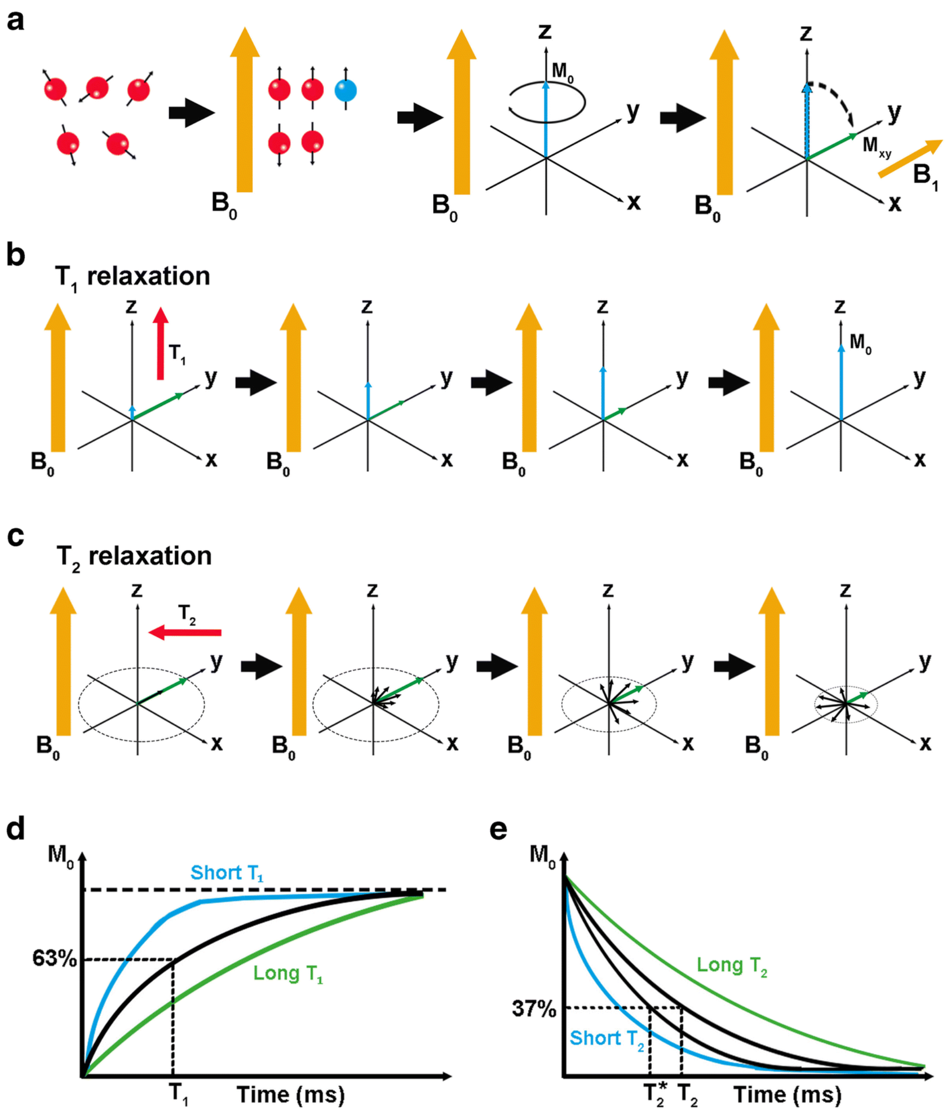


Figure 2.1: Figure from Mastrogiacomo et al., 2019. (a) At first the magnetic moments of the individual nuclei are oriented randomly and cancel out each other. In the presence of an external field B_0 they gain a net magnetization. With the presence of the external RF pulse the phase of the net magnetization is changed and magnetic resonance is said to occur. (b) Describes process of T1 relaxation in which the net magnetization starts to align along the static magnetic field B_0. (c) Describes T2 relaxation or the dephasing of magnetization in the transverse plane. (d) Illustrates that the value of the T1 constant is the time required to achieve 63% of the longitudinal magnetization. (e) The time constant T2 is the time required for the magnetization to fall to 1/e of its value. In (d) and (e) the blue lines represent tissues containing fat while the green lines represent fluids such as CSF.

is due to the release of electromagnetic energy into the environment to attain thermal equilibrium. This release of electromagnetic energy is what forms the signal for the receiver coil. Bloch, 1946a introduced two time constants to measure the relaxation phase, T1 and T2.

The constant T1 measures the growth of the longitudinal component (M_z, Fig. 2.1.b) with T1 relaxation being termed as the process by which the net magnetization aligns itself along the direction of the original magnetic field. The T2 constant measures the decay of the transverse component of the magnetization (Fig. 2.1.c). The value of T2 reflects the time required for the magnetization to fall to $\frac{1}{e}$ or 37% of its original value, illustrated in Fig. 2.1.e. Since there can be inhomogeneities inside the MRI scanner, another constant $T2^*$ is measured as the T2 time observed while taking the recording to account for such effects. During the resonance phenomena, the magnetization has components in different directions which can be expressed in terms of the time constants and time lapse (t) of the experiment.

$$M_x(t) = M_o e^{\frac{-t}{T2}} \sin \omega t \tag{2.4}$$

$$M_y(t) = M_o e^{\frac{-t}{T2}} \cos \omega t \tag{2.5}$$

$$M_z(t) = M_o(1 - e^{\frac{-t}{T1}}) \tag{2.6}$$

Where M_x ,M_y, M_z are the magnetizations along the x,y and z directions and M_0 is the magnetization induced by the applied static magnetic field B_0. From Eq. (2.6) it is clear that the T1 constant measure the time required to gain the original magnetization, i.e. $t = T1 \implies M_z = M_0$ and $t = T2 \implies M_{xy} = \frac{M_0}{\cos \omega T2}$. The differences in T1 and T2 times of different tissues (Fig. 2.1.d and Fig. 2.1.e) are often used to generate a contrast for image formation.

Image contrasts are indirectly controlled using repetition time (TR) and echo time (TE) in a sequence. TR is defined as the time taken between successive excitations of the same region. TE measures time between excitation and signal measurement. Short TR and short TE generate T1 weighted images while long TR and long TE generate T2 weighted images.

Different types of brain tissues have different T1 and T2 values. Short T1 times are seen in fat molecules because of their complex structure. The intricacies of saturated molecules lead to flexions and roations which might occur at the Larmour frequency. This makes it easier for the magnetization to return to the initial state (longitudinal orientation). Longer T1 times are observed in comparatively freely diffusing mediums such as CSF. Such fluids appear dark on T1-weighted images due to longer T1 times while fat molecules appear bright with the same diffusion weighting. Fat molecules appear lighter on T2-weighted images due to short T2 time while the CSF appears bright since it has long T2 times.

2.1.1 Image formation

An MRI scan needs to represent the 3D nature of specimen under study. Inside an MRI scanner, the gradient coils are used to alter the magnetic field along the different spatial directions. This ensures that different slices of the biological structure to resonate at different frequencies. The signal from all the slices is then detected by receiver coils that recognize only the transverse magnetization. In neuroimaging studies this signal needs spatial encoding to transmit information about the millions of voxels at different spatial locations in the brain.

It is known that in order to form the 3D image the magnetic gradients are applied in the x,y,z directions. This formalism results in each voxel possessing a different Larmour frequency and the phenomenon can be termed as spatial encoding.

$$\omega = \gamma(B_0 + G(x, y, z)) \tag{2.7}$$

According to Eq. (2.7), it is evident that there is a direct relation between the gradient field and the Larmour frequency. Usually the gradient for the slice selection is applied along the z-direction. For better understanding a simpler case of 2D image reconstruction has been explained in the Fig. 2.2 where the phase encoding gradients are applied in a direction perpendicular to the frequency encoding gradients.

The RF coils in the MRI scanner detects a signal containing a mixture of frequencies specific for each slice and each phase encoding step. The distribution of frequencies is determined using a Fourier transform. This distribution is then used to fill what is called as the k-space. The 2D k-space in Fig. 2.2 is used to elucidate the components of the frequency. The intensity of a point in the space represents the contribution of the frequency (k_x, k_y) in the signal. For each combination of (k_x, k_y) in the k-space the scanner camera takes only one picture (one filter per voxel). It then estimates the actual intensities at different locations using an inverse Fourier transform. There is a one-to-one correspondence of the pixel in the k-space to the image space. Every pixel in the 2D k-space image maps to only one pixel in the reconstructed 2D image. However, it is not necessary that the locations in both the images are exactly the same. This mapping using a Fourier transform is then extrapolated in 3D in order to obtain the image of the whole organ such as the brain. Any weighting such as T1w or T2w can be given in order to generate the image after the transformation has taken place.

Another important concept for image formation is that of field of view (FOV).

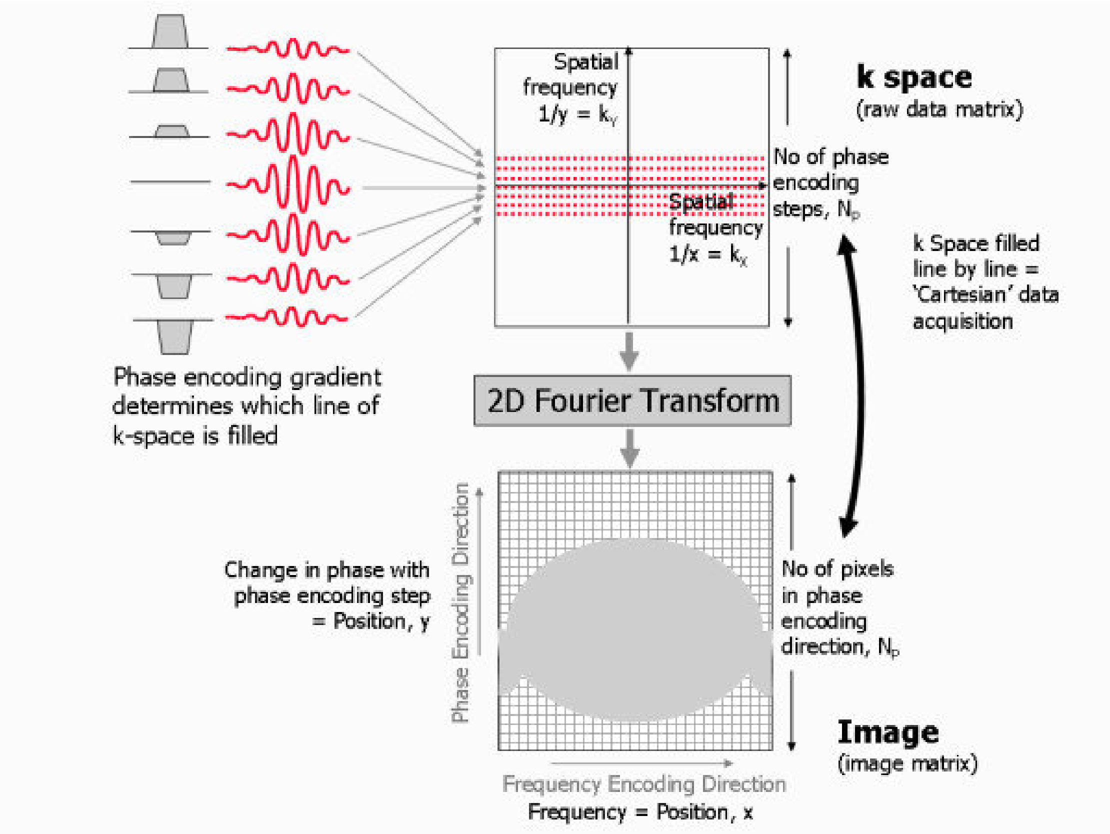


Figure 2.2: Image from Ridgway, 2010 summarising the reconstruction of an encoded MRI signal in 2D. A 2D slice has been selected using the gradient G_z. The phase is encoded along the y direction and the frequency is encoded along the x direction. Each phase encoding step is used to populate the k-space. The number of such steps determines the number of pixels along the y direction in the reconstructed image. The relationship between the k-space points and the points in reconstructed image space is such that $k_x = 1/x$ and $k_y = 1/y$. In the k-space, each line parallel to the k_x axis corresponds to a separate MRI signal. The number of such signals is the number of times the pulse sequence is repeated. The location on the x-axis determines the time during signal acquisition. Each line parallel to the k_y axis corresponds to the amplitude and duration of the phase encoding direction (gradient) at each phase encoding step (y coordinate).

It refers to the distance over which an MRI signal is acquired or displayed. The defined FOV determines the pixel width (determined by the phase encoding y direction in Fig. 2.2) given by $\Delta k = 1/FOV$.

2.1.2 Diffusion MRI

dMRI is an *in-vivo*, non-invasive imaging modality that used to create high-resolution structural images of biological tissues. It measures the non-homogeneity of water diffusion in tissues to probe their microstructure (Ghosh and Deriche, 2015). One of the most important applications of dMRI is to map the white-matter fiber tracts in the brain. As compared to other MRI techniques such as fMRI it does not suffer from an issue of low resolution and low signal-to-noise ratio (SNR) (Wong et al., 2016).

“Diffusion” is defined as the net movement of a substance from a region of higher concentration to a region of lower concentration. In a homogeneous medium, the diffusion of water molecules is isotropic, i.e. the molecules can move in any direction with equal probability. They exhibit random walk behavior which is explained by Brownian motion (Brogioli and Vailati, 2000).

The environment inside a biological tissue is complex and the diffusion of water molecules becomes anisotropic due to the hindrances imposed by cellular membranes. This causes water molecules in the extracellular environment to experience relatively free diffusion and the ones in the intra-cellular environment experience restricted diffusion (Toennies, n.d.). This diffusion anisotropy is encoded into the MRI signal using spatial and temporal variation (gradients) in the magnetic field. This means that the alignment of the molecules is influenced by the diffusion direction. The MRI signal is said to be “Diffusion Weighted” due to signal attenuation introduced by the magnetic field gradients.

2.1.2.1 Diffusion Weighted Imaging

In Diffusion Weighted Imaging (DWI) the intensity of each voxel represents the rate of water diffusion in a cubic region. Diffusion weighting is applied in order to generate contrasts based on the assumption that diffusion varies with pathology i.e. differences in diffusion can also highlight differences in structure and function (Taylor and Bushell, 1985).

One of the most popular ways to give images diffusion weighting is by using single-shot spin-echo (SE) T2 weighted sequences with two symmetric gradients on each side of the 180 degree refocusing pulse. This is based on the pulsed-gradient spin-echo (PGSE) technique developed by Stejskal and Tanner, 1965. PGSE im-

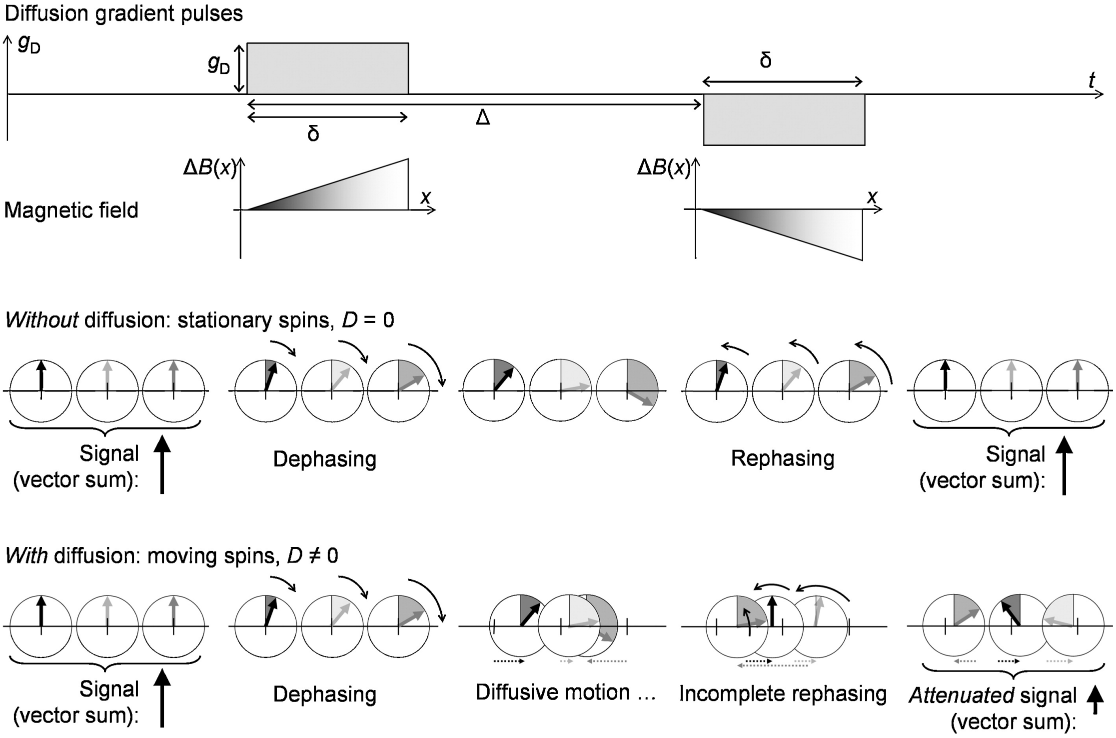


Figure 2.3: Image from Dietrich et al., 2010 elucidating the technical aspects of Diffusion Weighted MRI. The RF pulse causes a phase change of the net magnetization. Introducing the diffusion gradient adds spatial dependence of phase shift on the individual magnetic moments. The molecules that have restricted diffusion do not experience the effect of the diffusion gradient. Any phase change from the first gradient is reversed by the second. They tend to relax to their equilibrium state. However, the molecules that undergo free diffusion experience the effects of the diffusion gradients introduced by the second RF pulse (180 degrees) and hence will undergo a total phase shift dependent on the spatial location. This shift is manifested in terms of signal attenuation. The degree of signal attenuation depends on multiple factors as shown in Eq. (2.8).

proved sensitivity to diffusion in comparison to the steady state gradients used previously. Stejskal and Tanner, 1965 solved the Bloch-Torrey partial differential equations for a symmetric pair of pulsed gradients (Bloch, 1946b) and obtained the well-known Stejskal-Tanner formula:

$$S = S_0 \exp^{-bD} \tag{2.8}$$

Here S_0 represents the original signal strength, S is the signal strength in a pulse sequence with the presence of diffusion gradients (g_D) and D represents the diffusion coefficient. The attenuation of the dMRI signal by the diffusion gradients is represented in Fig. 2.3. The application of the diffusion gradient results in a spatial encoding as the Larmour frequency becomes dependent on the net magnetic field (Eq. (2.7)). In Fig. 2.3, it is evident that the Diffusion gradients are magnetic field gradients along the x direction.

Here $\Delta B_(x) = g_D x$, which makes the Larmour frequency spatially dependent $\Delta\omega = \gamma \Delta B_(x)$ or $\Delta\omega = \gamma g_D x$. There is a phase shift introduced after the gradient pulse of time duration δ. The the phase shift is also dependent on the x position due to the spatial dependence represented as $\Delta w(x)$, the phase shift is $\Delta\phi(x) = \Delta w \delta = \gamma g_D x \delta$. This spatial dependence of the phase shift makes spins at different positions along the gradient axis "dephased" after the application of the gradient pulse. When the negative gradient is applied, the process of rephasing occurs. The dephasing and rephasing mechanisms result in the diffusion weighting of the image. Without any diffusion stationary spins would align along their equilibrium position (cancelling effect of the two opposing diffusion gradients) while with diffusion weighting there is a signal attenuation explained by Eq. (2.8).

The b value of the diffusion weighted signal mentioned in Eq. (2.8) is defined in the units of s/mm^2 as

$$b = (\gamma g_D \delta)^2 (\Delta - \frac{\delta}{3}) \tag{2.9}$$

In order to obtain the numerical value of b, long and strong gradients are required. The diffusion gradients g_D, time of pulse δ and time interval Δ are often adjusted to adjust the b values. Higher b values leads to lower signal in the areas of high diffusion and increases the contrast between tissues that have different diffusion coefficients. The b values often need to be adjusted to obtain an optimal SNR.

2.1.2.2 Diffusion Tensor Model

DTI is a new type of imaging technique that relies on a tensor model to measure the diffusion on per voxel basis. Instead of attributing diffusion inside a voxel by to a single quantity, it uses a tensor formalism to measure diffusion along different directions within a voxel. The tensor model gives a rotationally invariant description of water diffusion. With this framework DTI is hence able to trace complex fiber tracts in the brain (D. Jones, 2010).

DTI is a novel technology, an *in-vivo* application of DWI and is the gold standard for imaging neural fiber tracts. It has become an important brain imaging modality since various neurological disorders such as cerebral ischemia and Parkinson's disease can be attributed to white matter defects. Further, white matter constitutes about 50% of the brain (by volume) which makes it important to understand both its structure and tissue composition. Analyzing structural brain connectivity using DTI does not only help to understand the patho-physiological effect of brain disorders but also the structure of its functional networks.

In this type of imaging, each voxel is associated with a 3×3 diffusion tensor representing the diffusion of water molecules using a Gaussian model. It is symmetric and contains six unique variables that characterize diffusion (as anisotropic or isotropic). This tensor has 3 eigenvalues and 3 corresponding eigenvectors which represent the directions of diffusion along the voxel. The voxels are usually $1mm^3$ in size and often constitute components of more than one cell within them. The diffusion tensor can be written as:

$$D = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix} \tag{2.10}$$

where $D_{xy} = D_{yx}$, $D_{zy} = D_{yz}$ and $D_{xz} = D_{zx}$ In this case Equation 2.8 can be written as

$$\frac{S}{S_0} = \exp^{-bg^T Dg} \tag{2.11}$$

where g^T is a 3×1 unit vector representing the gradient direction. The cellular environment is heterogeneous, so water molecules in certain parts undergo free diffusion while in others they undergo restricted diffusion. Due to this restricted diffusion the measured diffusion coefficient (of the water molecules) is different from the regular diffusion coefficient of water, it is termed as the apparent diffusion coefficient (ADC). Diffusion in complex environments cannot be explained by using diffusion gradients in one direction only. Figure 2.3 shows that only the component along the gradient direction is detected. Therefore, it is required to apply diffusion gradients in three directions to get an estimate of anisotropy of water molecules. Diffusion anisotropy in a voxel means the deviation of the voxel diffusion from isotropic diffusion. High diffusion anisotropy means that there is a preferred direction for the water molecules within that voxels to diffuse (Clark et al., 2011).

DTI images are usually represented by either encoding the tensor information using a scalar (for intensity values in a black and white image) or 4 numbers (red, green and blue (RGB) and brightness). Visually, the tensors can also be viewed as glyphs and very famously by tracing white matter tracts through a process known as tractography. The quantities such as mean diffusivity (MD) and ADC are used to characterize diffusion magnitude. Other scalars such as fractional anisotropy (FA) are used to characterize diffusion anisotropy.

MD is calculated as the trace of the diffusion tensor (Eq. 2.10). The FA determines an average ratio of diffusion distortion from the applied gradient directions. In order to calculate the FA, the diffusion tensor is converted to a diagonal matrix (which has eigenvalues D1, D2, D3 the diffusion coefficients along x, y and z directions).

$$D = \begin{pmatrix} D1 & 0 & 0 \\ 0 & D2 & 0 \\ 0 & 0 & D3 \end{pmatrix}$$

$$FA = \sqrt{\frac{3}{2}} \frac{\sqrt{(D1 - D_{mean})^2 + (D2 - D_{mean})^2 + (D3 - D_{mean})^2}}{\sqrt{D1^2 + D2^2 + D3^2}} \tag{2.12}$$

Where D1, D2, and D3 are the corresponding eigenvalues and v1,v2,v3 are the eigenvectors. DTI is a popular method to study the orientation and organisation of white matter. However, it fails in regions containing populations of fiber orientations that have different fiber orientations. It assumes that all white matter bundles in the brain have similar diffusion characteristics and attributes diffusion anisotropy to partial volume effects (J.-D. Tournier, Calamante, et al., 2004a). Secondly, the diffusion tensor only possess a single major eigenvalue for modelling the diffusion in one voxel and cannot be used for mixed fiber populations.

2.1.2.3 Higher Order Models

In order to solve the multiple fiber orientation problem of DTI mentioned in subsubsection 2.1.2.2, a number of approaches have been proposed to estimate the composition of fiber orientations inside a voxel. These models extract higher order structural information of tissues. They can help deal with problems such as kissing and crossing fibers.

High Angular Resolution Imaging (HARDI) techniques fall into the class of higher order models since they enable the detection of multi-modal diffusion signals. They include methods that incorporate the acquisition of diffusion data in more than 6 diffusion gradient directions. Q-ball imaging is one such technique but has its own limitations (J.-D. Tournier, Calamante, et al., 2004b). These methods rely on the concept of a fiber Orientation Distribution Function (fODF). An fODF is a symmetric probability distribution function describing the distribution of fiber orientations

$$F(\Theta,\phi)=\sum_{k=1}^{K} w_k \delta_{\theta_k,\phi_k}(\Theta,\phi) \qquad \Theta\in[0,\pi],\phi\in[0,2\pi] \tag{2.13}$$

where w_k represents the volume fraction of each fiber passing through the voxel. Θ_k and ϕ_k represent the polar and azimuthal angles in spherical coordinates respectively.

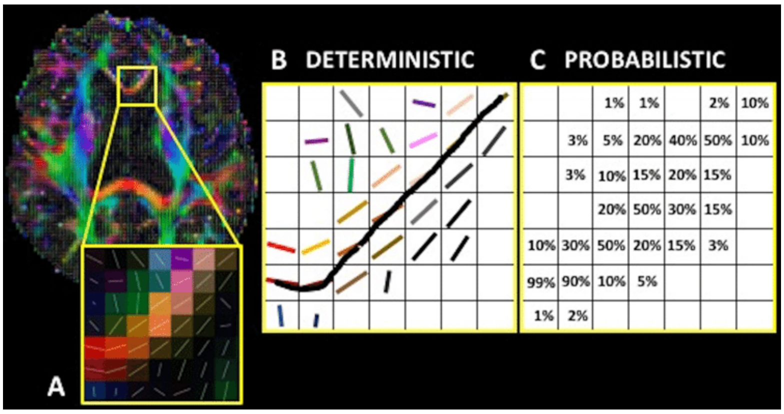
To overcome the limitations of methods such as q-ball imaging, J.-D. Tournier, Calamante, et al., 2004a propose modelling the DWI signal as the spherical convolution (i.e. convolution in spherical coordinates) of the response function and the fODF. A response function is mathematically an axially symmetric kernel which describes the DWI signal resulting from water diffusion along each fiber bundle aligned with the z-axis in a voxel. A different response function is estimated for different fiber populations such as white matter (WM), gray matter (GM) as CSF.

Spherical Deconvolution (SD) (Dell'Acqua and J.-D. Tournier, 2019) is a method in which the fODF can be extracted by deconvolving the response function from the DWI signal. Constrained Spherical Deconvolution (CSD) has also gained recent attention as a method to extract WM fiber orientation distributions as it imposes constraints to remove negative values in the reconstructed fODF. This increases the clinical feasibility and plausibility of the WM tractography (Jeurissen et al., 2014a).

2.1.2.4 Tractography

Fiber tractography is a technique used to graphically construct 3D representations of white matter pathways in the human brain. The technique uses the information about local fiber directions to trace streamlines that represent white matter tracts in the brain. It has become an important part of white matter disorder studies and has gained popularity because of its non-invasive nature. These advantages make it suitable for surgical targeting and planning (Romano et al., 2009).

Tractography methods often use DTI in order to delineate the streamlines along the white matter tracts. This is done by tracing the direction of diffusion from one voxel to the next. Fiber tracking relies on the assumption that the direction of maximum diffusion(in the diffusion tensor) aligns with the direction of the white matter fiber within the voxel. Tractography can be divided into two classes:



		1%	1%		2%	10%
	3%	5%	20%	40%	50%	10%
	3%	10%	15%	20%	15%	
		20%	50%	30%	15%	
10%	30%	50%	20%	15%	3%	
99%	90%	10%	5%			
1%	2%					

Figure 2.4: Image from Muller et al., 2018 depicting different types of fiber tracking techniques. Red, green and blue regions indicate tracts running along the x, y and z axis respectively. (a) Depicts deterministic tracking where the largest eigenvector determines the direction of progression along adjacent voxels. (b) Probabilistic tractography is depicted. A probability distribution function represents multiple possible diffusion directions.

deterministic and probabilistic tractography.

In deterministic tractography, the algorithm usually follows the direction of maximum diffusion from along the voxels by starting from a “seed” region. If the angle between two subsequent directions of maximum diffusion is less than predefined threshold, the algorithm proceeds tracking along the maximum diffusion direction of the second voxel and terminates when the condition is not met (Descoteaux et al., 2008). The deterministic tractography does not take into account any random or systematic errors that arise during signal acquisition, recording and transmission.

Probabilistic tractography relies on the estimation of a probability distribution function (PDF) such as an fODF for the distribution of fiber orientations in each voxel. This helps to account for the possibility of crossing fibers and kissing fibers. The PDFs and uncertainty are then used to build a path probability map corresponding to a seed point. Let’s say N streamlines are propagated randomly from a seed point using an orientation along the distribution functions. For any to voxels A and B

$$P_{AB} = \frac{M}{N}\text{ , probability that a curve starting at voxel A passes through voxel B}$$

$$M = \text{number of streamlines that go through B and A}$$

$$N = \text{total number of streamlines generated from A}$$

Using this probability map the most probable fiber orientation can be determined. From the multiple fiber orientations existing in a voxel the one most compatible with the incoming trajectory is chosen (T. E. Behrens et al., 2003).

2.2 Analyzing the Brain as a Graph

The human brain is one of the most intricate biological systems. It has been well established this complex system can be modeled as a network composed of various dynamically interacting elements. The networks perspective has gained tractable attention in neuroscience and given rise to an ever expanding field of network neuroscience. Brain organization can vary at scale i.e. from molecular interactions to cognition. With different interactions and emergent properties at each level of organization, graph theory has become an indispensable tool in this field (Sporns, 2018).

Brain networks can be effectively represented as graphs for computational analysis. In the graphical representations, each node could represent an entire brain region or an individual neuron. Edges can be binary or weighted, directed or undirected. The architecture and properties of the graph depend on the scale and the nature of interactions being considered (Rubinov and Sporns, 2010).

The widespread application of graph theory in neuroscience has given rise to the field of *connectomics* (Sporns, Tononi, et al., 2005) aimed at the analysis of the structural and functional brain connectivity. The structural networks usually represent the anatomy of the brain and are temporally stable (ignoring the effects of plasticity and development). On the other hand, the functional networks are temporally variable and dense. Structural networks are preferred when trying to study the physical nature of the brain and it's anatomy when trying to predict neurological disorders and searching the biochemical basis of cognition.

2.2.1 Connectomics

Connectomics is termed as the study of the brain's functional and structural networks. A connectome (in-vivo, retraced from imaging) is a dense network of brain connections, with a numerical value assigned to each network (Bassett and Sporns, 2017). The connectome of a brain can be seen as a circuit diagram with the neural connections analogous to wires and the cell bodies to electrical components.

Recently connectomics has become the focus of major neuroscience studies. Research in this field has expanded rapidly due to the increasing interest in understanding the brain as a dynamic system, and the belief that its connections give rise to its capabilities and functionality (Sporns and Bassett, 2018).

The first complete structural connectome to be mapped at the synaptic level was published in 1986 and belongs to the species C. elegans. It was reconstructed using electron micrographs of the serial section. It is a network that possesses three hundred neurons and roughly seven thousand connections. After 1986, it still took decades to claim the biological plausibility of all the connections in the reconstructed connectome (Cook et al., 2019). This trend lead to the conclusion that the connectome of the human brain cannot be mapped using the manual labor-intensive methodology employed by White et al., 1986. The human brain contains roughly the same number of neurons as the stars in the Milkyway galaxy and about 10^{15} inter-neuronal connections called synapses (Fornito et al., 2015). Mapping the human brain manually seems implausible in terms of time considerations. This calls for the need to computationally determine brain connectivity from brain scans.

dMRI is used for the construction of the structural connectome while fMRI is the preferred modality for functional connectome construction. The HCP was launched in 2007 as the first large scale collaborative effort to create detailed maps of the brain (in-vivo) and to help better understand the fundamentals of human connectional anatomy.

2.2.2 Structural Brain Connectivity

Structural brain connectivity refers to the arrangement of anatomical connections in the brain. A model of the brain's structural connectivity can be derived from whole brain tractography. In organisms with complex nervous systems such as that of humans, the structural brain connectivity can be visualized at different scales. It can be either at the level of synaptic connections (microscale), at the level of neuronal populations (meso-scale) or at macro scale in which fiber tracts run between different brain regions. At all these scales, the connectivity patterns of individuals from the same species exhibit different characteristics such as organization, topology and spatial extent (Sporns, 2007).

In neuroimaging studies the connectivity is usually analyzed at the macroscopic scale, i.e. fibers that run in between different brain regions are traced. From subsubsection 2.1.2.4 it can be inferred that the streamlines traced using the tractography algorithms represent the structural connections between different ROIs of the brain that are represented as brain nodes. The ROI-to-ROI connections are usually represented using bio-physical parameters such as mean FA, number of streamlines between the two nodes and the length of the streamlines connecting the two nodes.

2.3 Feature Selection Techniques

Neuroimaging data often suffers from the *curse of dimensionality.* The sample size is often much smaller than the total number of features. Classifiers might be influenced by the increase in noise as the number of features increases. Dimensionality reduction and feature selection are often the techniques used for a meaningful reduction in features. However, most dimensionality reduction methods such as Principal Component Analysis (PCA) rely on transformations of the original data and do not provide interpretable features. Due to this reason, it is important to consider feature selection methods where inference about which features the classifier learns can be made (Shi and Nathoo, 2018).

Feature selection methods can be grouped into three types: filter methods, wrapper methods and embedded methods. Filter methods are based on selecting features before running the classifier. Wrapper methods can be seen as a selection method which select features *on the fly* i.e. features are added or removed iteratively on the basis of classification performance. Embedded methods are those in which feature selection is embedded in the classifier. All three types of are being increasingly used in neuroimaging studies, however filter methods remain advantageous due to interpretability considerations (Tohka et al., 2016). The remainder of this section will focus on filter methods due to their relevance in the methodology of this thesis.

Filters form one of the simplest methods for feature selection. They usually serve as a preprocessing step for the classification and are model independent. However, there is a caveat in using such techniques for classification studies. The caveat is that each raw feature is considered independently and the cumulative effect of separate features is not taken into consideration for selection. One of the common techniques used to determine feature importances is the use of coefficients such as the f-scores, t-test and Pearson correlation coefficients which are widely used in neuroimaging applications (Mwangi et al., 2014).

Assume a two class classification task. In such a task the subjects can be divided into two groups. When the sample size is equal, a two independent samples t-test is a natural choice to compare the group differences. In such a case the null hypothesis remains that the means of the feature for the two groups is the same.

Based on Inza et al., 2004, the t-statistic is calculated as:

$$t = \frac{|\overline{x_1} - \overline{x_2}|}{\sigma_p * \sqrt{\frac{2}{n}}} \tag{2.14}$$

$$\sigma_p = \sqrt{(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})} \tag{2.15}$$

When the sample size of two groups are unequal. A Welsch's t-test can be carried out in cases where it can be assumed that the two groups have similar variances. The t-statistic is then given as:

$$t = \frac{|\overline{x_1} - \overline{x_2}|}{(\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}) * \sigma_p} \tag{2.16}$$

$$\sigma_p = \sqrt{\frac{(n_1 - 1) * \sigma_1^2 + (n_2 - 1) * \sigma_2^2}{n_1 + n_2 - 2}} \tag{2.17}$$

For both the t-statistics, σ_1 represents standard deviation for the feature values of the first group, σ_2 is the value for the second group, σ_p is the pooled standard deviation, and n_1, n_2 represent the number of samples for class 1 and 2 respectively. The features can be ranked on the basis of the p value of the t-test to obtain its statistical significance. Typically, values of $p < 0.05$ are considered statistically significant (Colquhoun, 2017) and only then can the null hypothesis can be rejected with confidence. Furthermore, using a t-statistic the population parameters can be estimated for a particular feature.

The t-tests have a number of limitations. One of the most prominent limitations is that can only be implemented for two group differences. F-scores can overcome this limitation and can be used to rank features for two or more groups. Mathematically, the f-score can be defined as:

$$F = \frac{\sum_{j=1}^{K}(\overline{x_j} - \overline{x})^2}{\frac{\sum_{j=1}^{K}\sum_{i=1}^{n1}(x_{j,i} - \overline{x_1})^2}{n_j - 1}} \tag{2.18}$$

where $\overline{x}$ represents the average value of the feature for all groups, $\overline{x_j}$ the feature average for for the $k_t h$ class.

The f-scores measures the ratio of between group variances to within group variances and the p-value of the t-test measures if the differences between the means of two groups are statistically significant. These two scoring techniques can be used to verify if the ranking generated by one corresponds to that generated by the other.

Based on a similar concept, the Pearson correlation coefficient can be used to filter features if the target variables are continuous. The Pearson correlation coefficient is defined as:

$$\rho = \frac{\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \overline{x})^2 \sum_{i=1}^{N}(y_i - \overline{y}^2)}} \tag{2.19}$$

where y represents the target variable and N represents the total number of data points. This takes the linear relationship between the feature values and the target variables into account.

2.4 Maximum Weight Connected Subgraph

Finding discriminative subgraphs representing interaction networks is an important problem in bioinformatics. One effective way of find such subgraphs is by solving the Maximum Weight Connected Subgraph (MWCS) problem. The major aim of the MWCS problem is to find a connected subgraph with the maximal sum of node weights (Loboda et al., 2016).

MWCS problem falls into the class of NP-hard problems. Given a connected and undirected graph $G = (V, E)$, the MWCS can be defined as that subgraph $G = (\tilde{V}, \tilde{E})$ which satisfies the equation:

$$\Omega(\tilde{G}) = \sum_{v \in \tilde{V}} w_v \longrightarrow max \tag{2.20}$$

where w_v is the weight of a node v.

There are many variants of this problem which are used for specific applications. For example, a cardinality-constrained MWCS is used in systems biology to detect core components in gene networks (Yamamoto et al., 2009). Analogously, a directed MWCS has been used to find the most deregulated networks in biological pathways. In both these cases the core networks are determined to be the ones which have maximum sum of the node weights (Backes et al., 2012).

The above stated implementations are not well suited for analysis of brain networks since connection strengths are of high importance in such a task. There has to be an account of similarity between nodes in order to get interpretable subgraphs. In such a case a generalized MWCS (GMWCS) based on the MWCS but with consideration of edge weights is suitable. The mathematical explanation for the generalized MWCS (GMWCS) will be presented in the remainder of this section.

Consider an undirected, connected graph $G = (V, E)$ with weighted nodes and edges. The weighting can either be positive or negative. Let V represent the set of its vertices, E the set of its nodes, w_e the weight of an edge e and w_v the weight of a node v. Mathematically, the generalized MWCS (GMWCS) can be described as that subgraph $\tilde{G} = (\tilde{V}, \tilde{E})$ which maximizes the following function:

$$\Omega(\tilde{G}) = \sum_{v \in \tilde{V}} w_v + \sum_{e \in \tilde{E}} w_e \longrightarrow max \tag{2.21}$$

Extracting such a subgraph can be a useful feature selection technique in connectomics. The usage of this method as a feature selection technique can help overcome the limitations of filter methods in neuroimaging. Filter methods do not incorporate spatial patters and interactions of multiple features (T. E. Behrens et al., 2003). Furthermore, there is a requirement to find a suitable thresholds for statistics such as p-value and t-statistic for which no ground truth is available. An implementation that finds a subgraph satisfying Eq. (2.21) eliminates the need to threshold and also takes the node interactions into account. Hence the GMWCS can be used as a feature selection technique in connectomics.

According to the conceptual framework in Loboda et al., 2016, the GMWCS problem can be formulated using a Mixed Integer Programming (MIP) formulation. By using a mixture of integral and binary variables, the maximal value of the objective function (Eq. 2.21) can be optimally found using a set of constraints.

The first step of formulating such a problem is to represent a candidate optimal subgraph. The membership of nodes and vertices in the subgraph is represented using binary variables:

$$y_v = 1 \iff v \in V \text{ and } v \in \tilde{V} \tag{2.22}$$

$$w_e = 1 \iff e \in E \text{ and } e \in \tilde{E} \tag{2.23}$$

Conceptually, an edge should be present in the subgraph only if both the end vertices are also present in the subgraph. In Eq. (2.24) this condition is represented as an inequality.

$$w_e \leq y_v \qquad \forall v \in V, e \in \delta_v \tag{2.24}$$

where δ_v represents the set of incoming edges on a node. Consider if $y_v = 0$, then w_e has to be zero because an edge cannot be present without the node being present and if $y_v = 1$ then $w_e = 0$ or $w_e = 1$ because it is not necessary for an edge to be present if a node is present.

After the subgraph can be represented, it is important to establish the validity of a candidate subgraph. A valid GMWCS needs to be connected and maximise the function Eq. (2.21). To establish the connectedness of the subgraph the concept of an arborescence needs to be introduced.

An arborescence is an acyclic directed graph in with a root node v, having exactly one path from v to any other vertex u. The connectedness in a graph is defined as the existence of a simple path between any two nodes. For creating a valid connected subgraph, its traversal has to be a valid arborescence. The GMWCS is constructed by finding a valid arborescence matching to its traversal.

The traversal of any graph can be found by considering $S = (V, A)$ derived from $G = (V, E)$ where each undirected edge (v, u) is replaced by two directed arcs (u, v) and (v, u). To ensure the connectedness of the subgraph, a non-linear formulation is needed where:

- Binary variable $x_a = 1$ iff $a \in A$ belongs to the arborescence.
- Binary variable $r_v = 1$ iff $v \in V$ is the root of the arborescence,
- Continuous variable $d_v = n$ if the length of the simple path from the root to the vertex v is n. The d_v value can be arbitrary if the vertex v does not belong to the optimal solution

As pointed above, a valid arborescence corresponds to the traversal of the connected subgraph. Hence, the arborescence needs satisfy the following constraints, according to Haouari et al., 2013:

$$\sum_{v \in V} r_v = 1 \tag{2.25}$$

$$1 \leq d_v \leq n \qquad \forall v \in V \tag{2.26}$$

$$\sum_{(u,v) \in A} x_{uv} + r_v = y_v \qquad \forall v \in V \tag{2.27}$$

$$x_{uv} + x_{vu} \leq w_e \qquad \forall e = (v, u) \in A \tag{2.28}$$

$$d_v r_v = r_v \tag{2.29}$$

$$d_u x_{vu} = (d_v + 1) x_{vu} \qquad \forall e = (v, u) \in A \tag{2.30}$$

Equation (2.25) ensures that there is only one root of the arborescence. Equation (2.26) limits the path length from the root vertex to any other vertex in the arborescence if it is present in the subgraph. With Eq. (2.27) it can be established if a vertex is a root then there are no incoming edges and that there is a maximum of one incoming edge on each node. The constraint in Eq. (2.28) says that either the forward arc or the backward arc can be a part of the valid arborescence and not both. The last two constraints in Eqs. (2.29) to (2.30) impose that the length of the path to the root is always 1 and the distance of the nodes correspond to the direction of the edges. It can be seen that the Eqs. (2.29) to (2.30) are non-linear. Combining the above equations the linearization is as follows:

$$d_v + nr_v \leq n \qquad \forall v \in V \tag{2.31}$$

$$n + d_u - d_v \geq (n+1)x_{vu} \qquad \forall (v,u) \in A \tag{2.32}$$

$$n + d_v - d_u \geq (n-1)x_{vu} \qquad \forall (v,u) \in A \tag{2.33}$$

Equation (2.31) combines Eqs. (2.25) to (2.26). Equations (2.32) to (2.33) can be justified using the fact that if a directed edge (v, u) is present, then node v is visited before node u and $d_u - d_v = 1$. They hence represent the Eq. (2.30). Any arbitrary node is considered to be the root of such an arborescence. The valid arborescence is then used to construct the optimal MEWIS.

2.5 Classification

An important task in connectomics is to analyze the differences between individuals using connectivity graphs. Machine learning algorithms have previously been used in combination with graphical representations in order to make predictions from neuroimaging data (Casanova et al., 2012). This section describes three classifiers which are incorporated in the methodology of this work.

2.5.1 Support Vector Classifiers

SVC are classical machine learning algorithms based on SVMs. They are important since they have attained significant accuracy with different types of tasks such as a handwritten digits recognition, face detection in images, and text categorization (C. Burges and C. J. Burges, 1998). In neuroimaging studies, SVCs are important for classification tasks since they are relatively robust to overfitting and more interpretable than Deep Learning classifiers.

In the simplest, binary case the mathematical formulation of the SVM is as follows. Consider each observation x_i, $i \in 1...n$ to be a vector in the *d-dimensional* feature space with a target label $y_i \in [-1, 1]$. The classifier needs to find a boundary that separates n such points present in the presented data within a small margin of error. It does so by trying to find an optimally separating hyperplane that efficiently divides the input data according to the target labels. In Fig. 2.5 the red line represents the optimally separating hyperplane that satisfies the equation

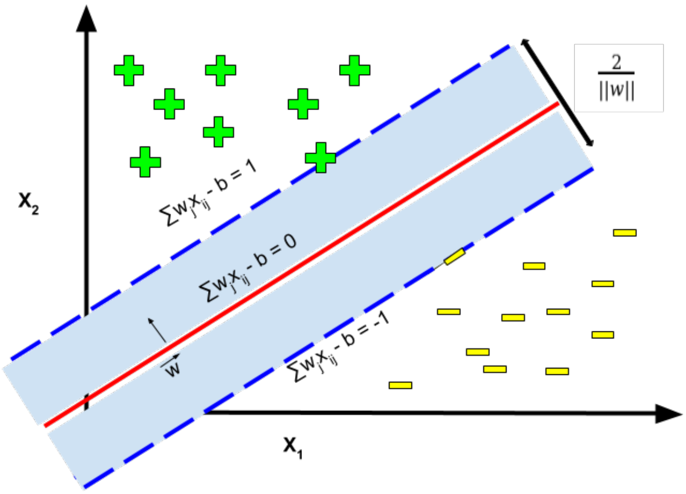


Figure 2.5: Illustration of classification using SVMs. The red line represents the decision boundary between the two classes. The algorithm maximizes the margin between the boundary for separate classes. The vectors lying on the hyperplanes depicted by the blue dashed lines are called as 'support vectors' and are the closest vectors to the decision boundary. Based on the values obtained by solving the left hand side of the decision boundary for each data points, they are assigned to the corresponding classes.

$\sum w_i x_i - b = 0$ It is maximally distant to the nearest point belonging to either class (also termed as the support vectors). The maximally separating hyperplane is found by satisfying the following constraints for each data point i.

$$y_i(\sum_{j=1}^{d} w_j x_{ij} + b) - 1 \geq 0 \qquad \forall i \in 1...n \tag{2.34}$$

Here, the weights represent the parameters the model learns in order to satisfy the above conditions by solving a Langrangian equation, the mathematics of which is beyond the scope of this text.

In cases where the decision boundary is a non-linear function of the data the algorithm makes uses of what is commonly called as the 'kernel trick'. In this method the pairwise dot products of the individual x_i are replaced by a non-linear transformation or kernel function. This expression then allows the algorithm to fit the maximum margin hyperplane in the transformed space. The decision boundary is linear in the transformed space and can be projected back to find the non-linear decision boundary in the original d dimensional feature space.

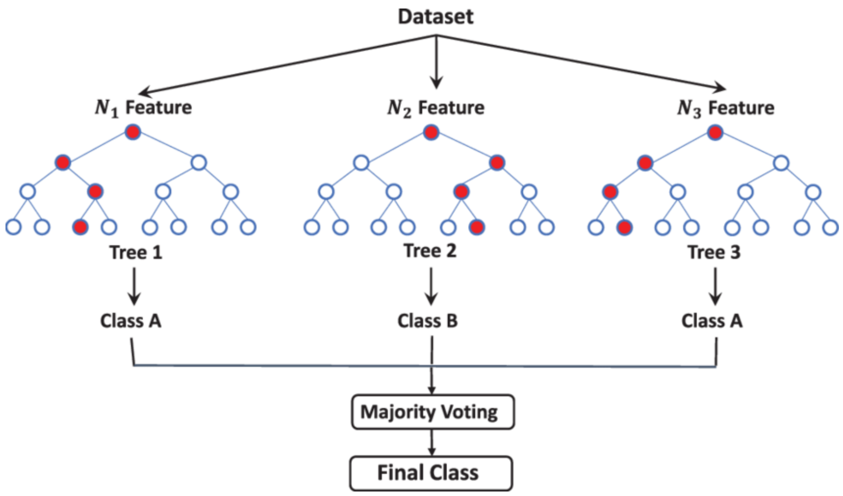


Figure 2.6: Schematic explanation of Random Forest classifier from Tahmasebi et al., 2020. The different trees represent the decision trees constructed from the bootstrapped samples. Each tree is created on the basis of the most discriminatory features, N_x for a particular bootstrapped sample. The final class of the samples in the dataset is decided on the basis of majority voting.

2.5.2 Random Forest Classifiers

RF Classifiers are based on the idea of bagging or bootstrap aggregation of decision trees (Hastie et al., 2009). A decision tree is a way of recursively splitting the target variables on the basis of the rules set on the features. The name 'Random Forest' comes from the fact that the algorithm builds a 'forest' by aggregating a large number of de-correlated trees and averages them for building a classification.

The Random Forest is built in the following manner. A number of runs is specified. First, a bootstrap sample is drawn from training data. On this bootstrapped data, one tree is grown by recursively splitting the tree tree until the minimum node size n_{min} is reached. The steps for the recursive decision tree construction are:

- Select k variables randomly from the p total variables.
- Find the most predictive variable that has the most discriminatory split point.
- Split the node into daughter nodes.

This method is repeated to obtain a decision tree for each time a bootstrapped sample is obtained. The individual trees are the aggregated to make an ensemble on the basis of the ensemble vote of each tree. It is evident from the Fig. 2.6, the

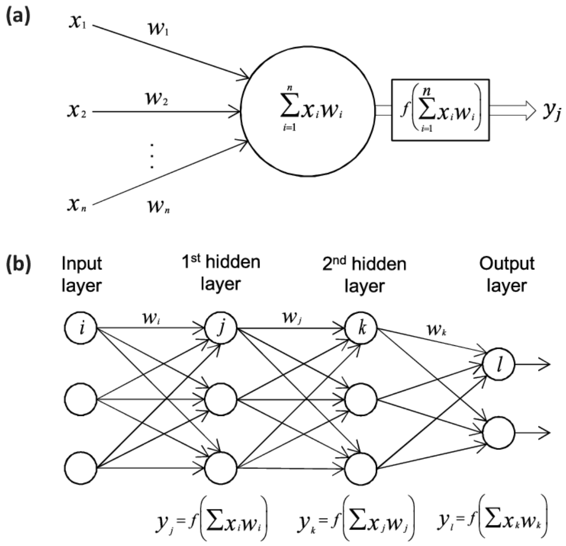


Figure 2.7: (a) Visualization of an artificial neuron from Vieira et al., 2017 depicting the output as a function of weighted sum of the inputs (Eq. 2.35) (b). The artificial neural network consisting of the input layer, hidden layers and the output layer.

final classification of samples is based on the majority voting and hence a feature importance can be determined. The higher the position of a split in the random forest tree, the higher is its discriminatory power. Random forest classifiers are widely used in neuroimaging studies due to the interpretability of features which can be obtained using the feature ranking in terms of the feature importance.

2.5.3 Multilayer Perceptron

A MLP is an Artificial Neural Network (ANN) that is organized in the form of layers to mimic biological neural networks. The network consists of an input layer, an output layer as well as one or more hidden layers as shown in Fig. 2.7.b. Each layer consists of one or more artificial neurons (also called as perceptrons) which are connected to neurons of the subsequent layers.

The connections between the layers are feed-forward and uni-directional. The input layer serves as a buffer layer with no transformation of the input while in the other layers the neurons implement non-linear transfer functions on the weighted connections from the previous layer as shown in Fig. 2.7.a. The output y from a neuron can be expressed as:

$$y = g(\sum_{i=1}^{N} w_i x_i + b) \tag{2.35}$$

where w_i represents the weighting of the inputs and b represents the bias for the neuron. Using the organizational structure, the neural network is able to learn complex transformations from the input data. In fact, it has been shown that an MLP with just one hidden layer and a finite number of neurons is able to act as a universal function approximator (Sifaoui et al., 2008). Consider that we have an input vector in the N dimensional space and the output is needed in the M dimensional space. An MLP (having non-linear transfer functional units) can implement any continuous mapping from the M dimensional space to the N dimensional space, with arbitrary accuracy. However, the number of neurons in the hidden layer required to achieve sufficient accuracy might be large which makes such a network difficult to train.

In order to achieve better accuracy and learn hierarchical representations, the network can be made 'deep' by adding more layers. Additionally, increasing the depth of the network helps overcome the problem of overfitting in shallow neural network without the need to train on a large number of samples (Bengio and Delalleau, 2011). Deep Learning is employed in neuroimaging studies in use-cases where there is a requirement to eliminate the need for manual feature selection.

3 Methods

A classification task was designed for the MRI data from the HCP. The pipeline for designing this classification task consisted of four major steps. First a whole brain tractography was generated on the basis of DWI scans for each subject. Then, three types of connectomes were created for each subject and represented in the form of connectivity matrices. This was followed by deployment of two different feature selection techniques for classification of subjects based on the connectivity matrices: namely 'baseline' and MEWIS. Finally, machine learning classifiers were used to test the effectiveness of these two feature selection techniques. The summary of the pipeline is presented in Figure 3.1.

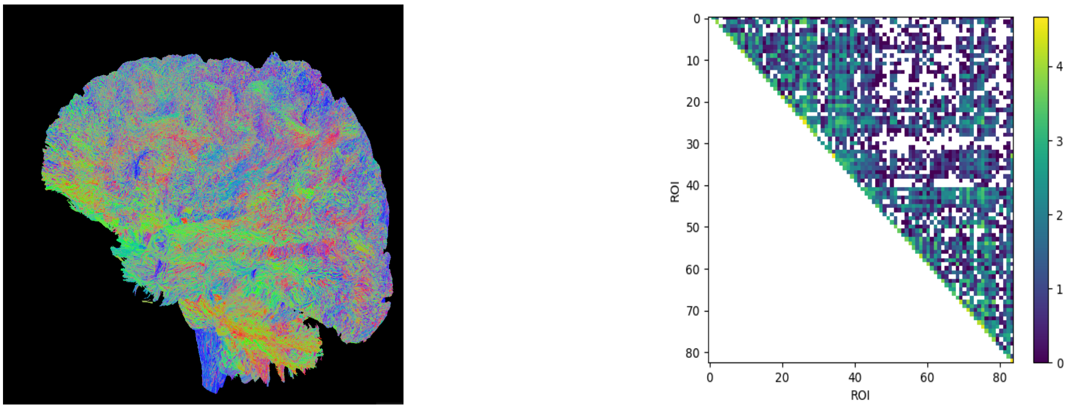
The implementation of this pipeline was scripted in *Python*. The preprocessing of the HCP data to generate tractography was done with the help of *MRtrix3* (J.-D. Tournier, R. Smith, et al., 2019). Data was prepared in a classification ready format using *Pandas* (McKinney et al., 2011) dataframes. The Maximum Edge Weight k-Induced Subgraph problem was implemented in Java based on a modification from Loboda et al., 2016 (***github.com/ctlab/gmwcs-solver***). The final classification was based on machine learning implementation from *scikit-learn* (Pedregosa et al., 2011).

3.1 Data Acquisition

Structural and dMRI data for 203 subjects was acquired from the s900 release of the HCP (HCP, 2015). Out of the total, 101 subjects were females and 102 were males. 83 females and 58 males were aged 26-30 and remaining 34 males and 28 females were aged 22-25. The demographic information along with other markers of emotion and cognition were obtained from the unrestricted access data available on ***https://db.humanconnectome.org/***.

3.1.1 Imaging Data

The structural and dMRI files used for this project were obtained from the repositories containing images processed using version 3 of the preprocessing pipelines in the HCP detailed in Glasser et al., 2013. A Siemens 3T Skyra system was used



(a) Tractography **(b)** Connectome

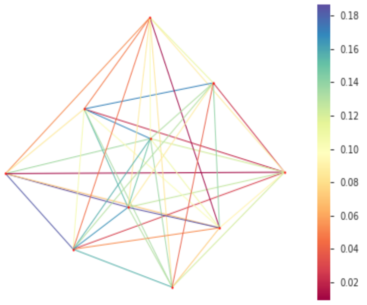
Feature Representation

f-scores, t-test, Pearson correlation

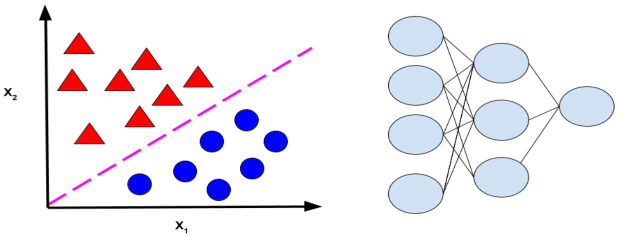
(c)

Feature Selection

Select k% features



(d) Feature selection using either baseline analysis or MEWIS solver.



(e) Classification

Figure 3.1: Summary of the pipeline implemented in this thesis. (a) Whole brain, one million streamlines tractography computed for each subject. (b) Connectivity matrix encoding DWI information. In this case, the matrix represents the number of streamlines between any two regions of interest in the log_{10} scale. (c) Different statistical measures used to represent the importance of the original features present in the dataset. (d) Ranking of edge importances using two different techniques. Baseline selects according to percentile distribution and MEWIS solver extracts a predictive subgraph. (e) Classification using SVMs, RFs and MLPs.

TR (ms)	2400
TE (ms)	2.14
T1 (ms)	1000
Flip angle	8 deg
FOV	224 x 224
Voxel Size	0.77 mm isotropic

Table 3.1: Acquisition parameters for the structural image acquired from the HCP s900 release.

Sequence	Spin-echo EPI
slice thickness	1.25 mm, 1.25 mm isotropic voxels
TR (ms)	5520
TE (ms)	89.5
Flip angle	78 deg
Refocusing flip angle	180 deg
FOV	224 x 224
Voxel Size	0.77 mm isotropic
b-values	100,2000 and 3000 s/mm^2

Table 3.2: Parameters for the acquisition of Diffusion MRI data acquired from the HCP.

used to scan all subjects (starting in August 2012, housed at Washington University, St. Louis). The details of the acquisition protocol are can be obtained from Van Essen, Ugurbil, et al., 2012.

In order to accomplish the goals of this project, there were two types of structural images acquired for each subject from the HCP pipeline. The first type was an image consisting of a segmentation volume along with cortical surface parcellation based on the Desikan Killiany Atlas (Desikan et al., 2006). The second type was the T1w scan in subject space. This anatomical images was aligned to the location at the interface of hemispheres: anterior and posterior commisure (AC-PC). This AC-PC alignment is a rigid body rotation so that the image gets 'centered' and can be termed as registration. The image was sampled at the same resolution as the diffusion data (1.25 mm isotropic, originally 0.77 mm isotropic). The parameters of the T1w images are presented in Table 3.1.

The important characteristics of the dMRI images acquired is that they were obtained in very high resolution (1.25 mm isotropic) using a Stejskal-Tanner (monopolar) diffusion encoding scheme as mentioned in Figure 2.1.2.1. The q-space was sampled by including 3 shells at the b-values presented in Table 3.2 with each gradient table defined by a single b-value acquired once with right-to-left and another in the opposite phase encoding polarities. For each subject, four types of DWI files were acquired: preprocessed diffusion time series, the brain mask in diffusion space, diffusion weighting for each volume and diffusion direction for each volume.

3.1.2 Label Preparation

	Agreeableness	**Openness**	**Conscientiousness**	**Neuroticism**	**Extraversion**
count	141.0	141.0	141.0	141.0	141.0
mean	32.482	28.326	35.262	16.965	30.766
std	5.144	6.021	5.615	8.046	5.918
min	13.0	12.0	21.0	0.0	18.0
25%	30.0	25.0	31.0	12.0	27.0
50%	33.0	28.0	36.0	17.0	31.0
75%	36.0	32.0	39.0	21.0	34.0
max	44.0	43.0	47.0	43.0	46.0

Table 3.3: Summary of personality traits for training data subjects.

The generalized pipeline implemented in this work could be used with different types of variables. Results for the big5 personality traits and gender are presented in the results section. Each of these labels were dealt with in a different manner.

The big5 personality traits are continuous. The statistical description of the personality traits for the training subjects is represented in Table 3.3. Personality traits are based on the five factor model of personality (Costa Jr and McCrae, 1992). The five different personality traits are agreeableness, conscientiousness, neuroticism, extraversion and openness. They are often used in psychology and psychiatry to characterize behavior.

The architecture of the pipeline of this work required the conversion of the continuous label prediction to a classification task. The conversion steps were as follows:

1. Record the median of the personality trait the training subjects.

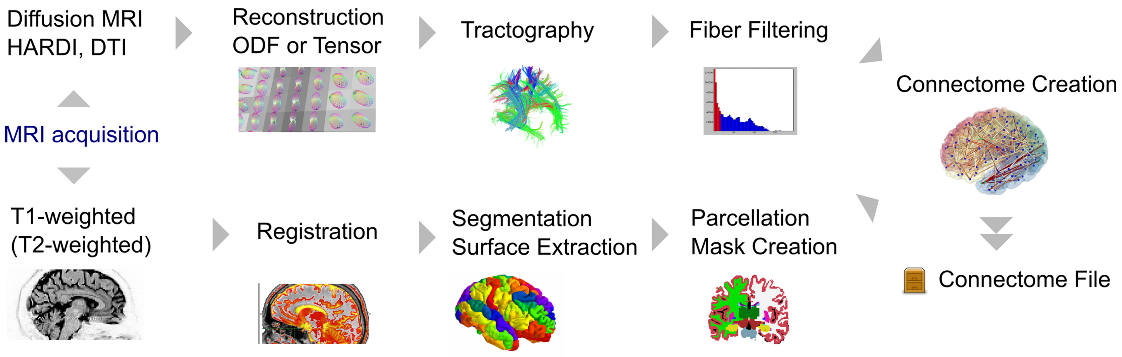


Figure 3.2: Pipeline for creating the connectome for each subject. Two parallel workflows following data acquisition have been illustrated in this figure. The first procedure aims to extract a parcellation mask from the structural images of each subject in their native space. The second 'parallel' procedure generates anatomically correspondent tractograms from diffusion images. A connectome file is generated after combining the information from filtered tractograms and parcellation mask. The connectome file contains properties of the fibers that connect ROIs determined by the parcellation mask. Image from Gerhard et al., 2011.

2. Divide the personality labels into 5 quartiles.

3. Remove the data of the subjects whose personality traits fall into the middle quartiles.

4. Binarize the variables such that the first two quartiles belong to the lower class and the last two quartiles correspond to the higher class.

The gender variables were categorical. For the classification task, male attribute was mapped to zero and females to one.

3.2 Creating the Connectome

A connectome for each subject was generated on the basis of tractography. The nodes of the connectome were determined using a parcellation mask as elaborated in subsection 2.2.1. This workflow was implemented on the basis of the tutorial titled "*ISMRM tutorial - Structural connectome for HCP*" . The preparation of the structural connectivity matrices has been visualized in Fig. 3.2. The subsquent subsections will explain the two parallel procedures for structural and diffusion images respectively which are combined to generate the connectome file.

3.2.1 Structural Image Processing

The end goal of the structural image processing was to generate a parcellation mask indicating gray matter ROIs. The steps of the structural image processing in Fig. 3.2 were readily accomplished by the HCP preprocessing pipeline (Glasser et al., 2013).

The registered T1w images along with the segmentation surfaces and parcellation masks were available within the HCP data (section 3.1). The T1w images were registered according to the rigid transformation for making the image centered along anterior-posterior commisure. The segmentation surfaces were labelled according to FreeSurfer's default subcortical segmentation consisting of 9 ROIs . The parcellation was based on FreeSurfer's automatic cortical parcellation according to the Desikan-Killiany atlas that delineates 35 ROIs. The combined parcellation and segmentation image was available from the HCP data was used to generate a volume delineating locations of the nodes of the connectome or ROIs.

FreeSurfer's default lookup table for the labels of ROIs was not linear. The nodes of the original 'combined' image were then mapped to the default lookup table in *MRtrix3*. This made the numbering of the ROIs (also the nodes of the connectome) start from 1. Furthermore, FreeSufer's estimates of sub-cortical grey matter structures were replaced with the estimates from FSL's FIRST tool. The nodes of the generated connectome were labelled according to the default lookup table in *MRtrix3*. There were 8 subcortical gray matter regions (the accumbens area was incorporated into thalamus label) and 34 cortical gray matter ROIs (see Appendix).

It is important to note that along with the diffusion data, the structural data was also used for probabilistic whole-brain tractography (Parker et al., 2003). Hence, in addition to generation of a combined parcellation mask, a five-tissue type (5TT) segmented image was also generated. This was done on the basis of sampling the T1w structural volume in the subject's native space at the same resolution as the diffusion data (1.25 mm isotropic). The 5TT image delineated the segmentation of brain regions into five tissue types; namely gray matter, subcortical gray matter, WM, CSF and pathological tissue. The segmentation is based on FSL segmentation tools FIRST and FAST. This information about the location of different tissue types makes the tracking based on DWI images suitable for ACT (R. E. Smith et al., 2012).

3.2.2 Diffusion Image Processing

The first step of processing the diffusion images was to determine the local fiber directions in order to prepare for probabilistic tractography. In order to accomplish this task the response functions of the WM, GM and CSF were extracted from the five-tissue type (5TT) image. From the Contrained Spherical Deconvolution (CSD) of the response function and the DWI image the fODFs were reconstructed based on the algorithm in Jeurissen et al., 2014b.

The probabilistic whole-brain tractography of five-million fiber tracts was generated using the white matter fiber orientation distributions. Tractography was performed in the subject space since a structural image in this space is the best approximation of the subject's physical brain. The streamlines were computed on the basis of the algorithm iFOD2 (J. D. Tournier et al., 2010). This algorithm uses the fODF image and determines candidate streamline paths (arcs), which have greater fODF amplitudes along the paths. By sampling the underlying fODF amplitudes along these arcs it makes the streamlines likely to follow most probable paths.

Anatomical constraints on the tractography were provided using the 5TT image. There were other series of constraints imposed in order to make the tractography more informed. The seed points were dynamically determined using the Spherical-deconvolution Informed Filtering of Tractograms (SIFT) model (R. E. Smith et al., 2013) of the white matter fODFs. The cutoff value of 0.06 was set for the probability amplitude for terminating tracks. The maximum length of streamlines was set at 250 mm i.e. 200 times the voxel size in our case where the voxel size is 1.25mm. The tracking along a streamline was truncated if the fiber terminates at poor structural location and retracking was performed. The streamlines were cropped whenever they cross the grey matter-white matter interface.

The tractography was downsampled from five million fibers to one million fibers to preserve the most biologically relevant fibers using the SIFT algorithm (R. E. Smith et al., 2013). This provided more meaningful estimates of the structural connection density and also reduced the memory requirements.

3.2.3 Connectome Representation

The connectome can be represented in the form of a graph as mentioned in subsection 2.2.1. After obtaining the tractography and a parcellation image representing the subject specific node locations of the connectome, the connectome matrix was obtained in a *.csv* format. It was generated from the whole brain tractography of one million tracts using 84 relevant grey matter parcellations (see Appendix).

The 84 nodes represent a 8 subcortical gray matter ROIs and 34 cortical gray matter ROIs with each ROI contributing to two nodes (one in the left hemisphere and another in the right hemisphere). Using different parameter settings three types of features were extracted. The mean FA (Eq. 2.12), the average length and number of streamlines between any two regions. The visualization can be seen in the Figure 3.1b. The connectome is represented as an upper triangular matrix (84×84, corresponding to 3570 features) considering that the connections between two ROIs are symmetric.

3.3 Feature Representation

Once the connectivity matrices for all the 201 subjects were computed and encoded in the format of a .csv file, a *pandas dataframe* was prepared to serve as input to the classifier. Each row represented the data for an individual subject. Each column represented the connection between any two brain regions i.e. features in one cell of the connectivity matrix. Three different types of features obtained using the connectome files according to subsection 3.2.3. The design of the classification task was based on different attributes presented in Figure 3.3. A classifier was trained on one set of configurations for two different feature selection techniques. The details will be discussed in the subsequent subsections.

Self-loops were omitted from the analysis pipeline due to incompatibility with the solver based implementation and non-significant changes to classification accuracy based on the omission. Furthermore, using the calculation of statistical coefficients on the raw data the importance of each feature was determined.

The training set consisted of 141 subjects consisting of 83 females and 58 males aged 26-30. While the independent test set consisted of 34 male and 28 females aged 22-25. The data was standardized using *sklearn* by removing the mean and scaling to unit variance since the different types of features were in separate scales of magnitude.

3.3.1 Statistical Coefficients

The raw features were not representative of any information on a group level. To gain information about feature importance statistical coefficients were used. From the raw matrices, each column was be represented as a statistic and hence ‘group averaged’ edge of the connectivity graph.

Statistical coefficients are effective feature filter methods for the classification from Neuroimaging data. Filter methods were used since they can be retraced to the original features. Mainly, three coefficients were used. The first two were

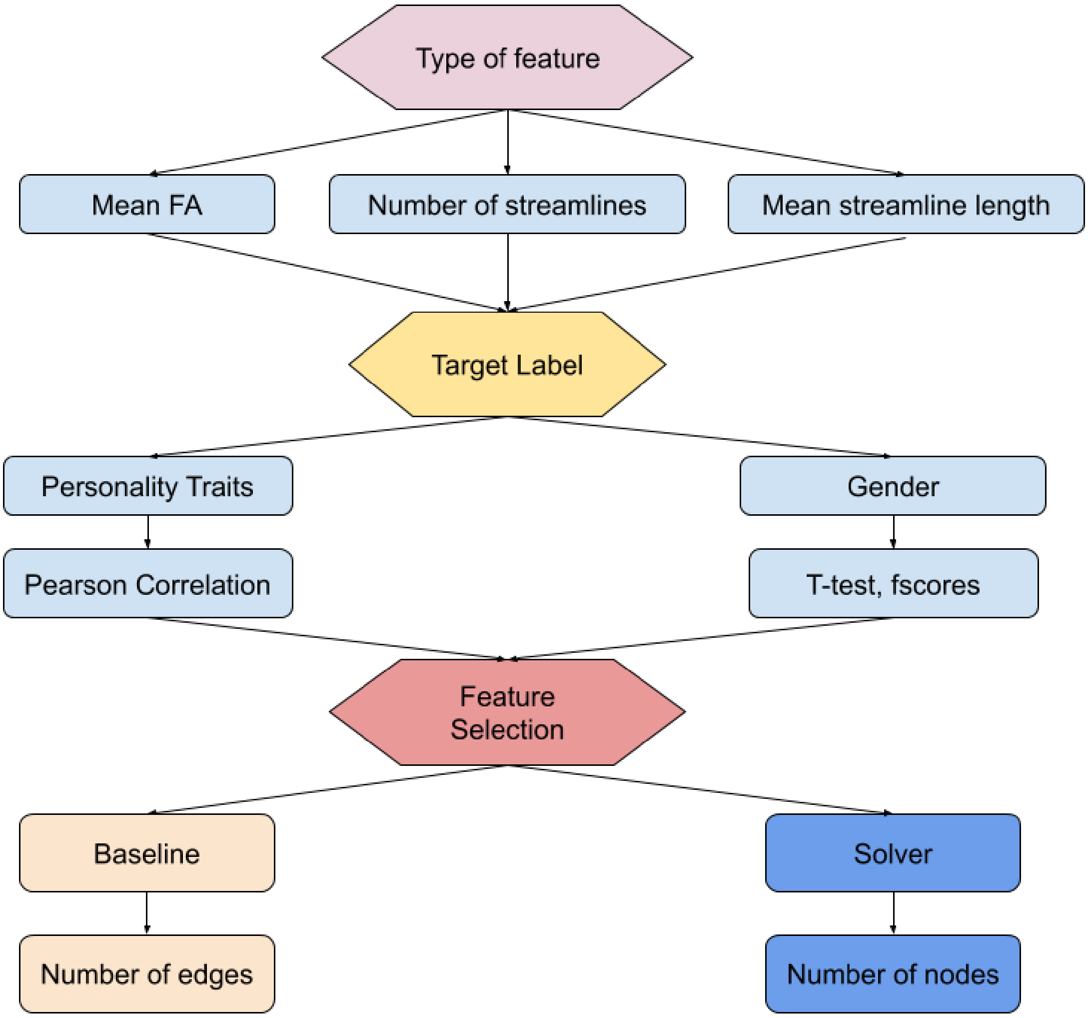


Figure 3.3: Considerations for classification task. Each path along this graph represents a different possibility for a classifier to be trained on. The parameters for training the classifier are: the type of feature, the target label and the feature selection technique.

the t-test and the f-scores,which describe the discriminative power of a particular feature. The Pearson correlation coefficient was used in the case of continuous variables to capture the linear relationships between the features and the target values. In order to rank the features an absolute value of the Pearson coefficient was used to highlight its importance.The f-score used in the analysis is used to measure how well the particular feature distinguishes between the two classes labelled as 1 and 2. It was calculated according to Eq. (2.18).

The t-test was processed as differently depending on the size of the classes. In the case of personality traits, size of both the classes was the same since the personality traits were binned according to the median (subsection 3.1.2). An independent samples t-test according to Eq. (2.14) was carried out for feature selection. For gender classification the sample size for both classes was different, the training set consisted of 83 females and 58 males. In this case an Welsch's t-test assuming similar variance was carried out (Eq. 2.16). The p-value of the t-test was used to determine it's statistical significance by converting it into the log_{10} scale. Each p-value p_x was represented by $f(p_x) = (-1) \times \log_{10} p_x$ so that a higher numerical value represents a higher statistical significance.

3.3.2 Exclusion of Self Loops

Self loops can be termed as the connections from the brain regions to themselves. They were excluded from the analysis before feeding data to the classifier. These loops were incompatible with the MEWIS implementations and hence they were removed from baseline analysis performance comparison between the two methods. Further, this exclusion will be justified by the performance of the classifiers with and without the inclusion of self-connections between the ROIs in subsection 4.3.2.

Their exclusion was an important consideration for the MEWIS solver based implementation. Degenerate edges are not part of a simple graph, their inclusion was incompatible with the implementation from Loboda et al., 2016. After removing such connections the MEWIS implementation was able to solve the MIP formulation and produce subgraphs.

The results in subsection 4.3.2 help evaluate the importance of self loops was ascertained by seeing their effects on classification metrics. This is done on the basis of a paired samples t-test. The first sample is taken as the classification performance on the data including the self loop features. Alternatively, the second sample is the classification performance obtained by excluding self loops. In general, the null hypothesis of a paired sample t-test remains that the differences between the observations from two samples is zero. Two slightly different experiments were carried out for classification of personality traits and gender. The formulas for calculation however remain the same. The t-statistic is calculated by:

$$t = \frac{\tilde{X_D} - \mu_0}{\frac{s_D}{\sqrt{n}}} \tag{3.1}$$

$$\tilde{X_D} = \sum_{i=1}^{N} \frac{a_i - b_i}{n} \tag{3.2}$$

where s_D is the standard deviation of the differences between the metrics of the experiments with the self loops a and without the self loops b. Arrays a and b have the same length n. There was one entry in the array for each possibility in a combination of choice for different attributes:

- Classification label
- Percentage of features in the set: $[2, 5, 10, 50, 100]$
- Three different classifiers: MLP, support vector machines, and random forest classifier
- Edge type (representatives of raw features, subsection 3.3.1): depending on classification label

The null hypothesis for classification of personality traits was that the average of a particular classification metric on the basis of one feature (such as mean FA, mean streamline length and number of streamlines) is the same. Including data for all five personality traits and one type of edge for arrays a and b had a size $n = 75$ i.e. $5 \times 5 \times 3 \times 1$.

In a similar fashion, the null hypothesis for the gender based classification was the same. However, instead of five personality traits there was one label used i.e. gender and two edge type (t-test and f-scores). This resulted in $n = 30$ i.e. $1 \times 5 \times 3 \times 2$.

The null hypothesis could not be rejected because for the test data, the p-values of these t-tests were high. This indicates that the difference between the classification on the data with and without the self loops is not statistically significant. Self loops could hence be safely eliminated and the results of these experiment are presented in subsection 4.3.2. After the elimination of the self loops that correspond to the diagonal of 84×84 connectivity matrix, there was a total of $n = 3486$ features for each subject.

3.4 Feature Selection

The motivation to perform feature selection from considerations of time complexity and removal of redundant features. There were two types of feature selection techniques used before classification. The first one was a classical feature selection technique based on statistical coefficients and is termed as the ‘baseline’. The second technique was based on extracting a subgraph after converting the original edge weights to statistical coefficients and is termed as the ‘solver’ method based on the MEWIS.

The MEWIS technique produces more interpretable results as compared to filter methods due to incorporation of graph topology. It is proposed due to the fact that analyzing the inter-subject differences at the subnetwork level is easier than analyzing dense subgraph of whole brain connectivity.

3.4.1 Baseline Analysis

Baseline experiments were solely based on the filter methods mentioned in section 2.3. In this technique, there is no information about graph topology or the biological correspondence of the features. Feature importance is determined purely on numerical values.

For each of the structural connections, the statistical coefficient was calculated based on the data and target variables from the training set. The metrics mentioned in subsection 3.3.1 were then used to rank the features. This ranking was based on first taking the absolute value of the coefficients and then dividing the coefficients for all features into a percentile distribution. This gave a ranking from which the top percentiles were selected according to a parameter k. Initially, $k \in \{2, 5, 10, 50, 100\}$ were chosen to cover the overall trend of classifier performance as a function number of features. Later in the course of multiple experiments, this parameter k could be adjusted according to requirements for comparison to the number of features extracted by the MEWIS solver.

3.4.2 Maximum Edge Weight k-Induced Subgraph

The Maximum Edge Weight k-Induced Subgraph (MEWIS) problem is defined as the extraction of a subgraph induced by preserving k nodes with a maximum total edge weight. The MEWIS was taken as special case of the GMWCS (section 2.4) for selecting graph based features in this pipeline. The implementation for the MEWIS in this project was based on a modification of the Java application from Loboda et al., 2016 which made use of the IBM ILOG CPLEX Studio version

12.10 Java Application Processing Interface (API).The general explanation of the MEWIS is as follows:

Consider an undirected graph $G = (V, E)$, then the aim is to find an induced subgraph $\tilde{G}(\tilde{V}, \tilde{E})$ with the set of vertices $\tilde{V} \subset V$ such that $\|\tilde{V}\| = k$, $\tilde{G}$ is connected and has a maximized sum of edge weights $\sum_{v \in \tilde{V}} w_e \longrightarrow max$.

The generalized Maximum Weight Connected Subgraph (MCWS) was converted into the MEWIS problem for a constant number of nodes. This was done by applying a constant negative weighting to all the nodes so that the subgraph which maximizes the MEWIS condition can be found. Based on Eq. (2.21) the MEWIS condition was modelled as:

$$\Phi(\tilde{G}) = j \times m + \sum_{e \in \tilde{E}} w_e \longrightarrow max \tag{3.3}$$

where each node has a constant weight j and the number of nodes corresponds to m. Such a technique was used for feature selection from structural brain connectivity as there is no predefined ranking for different regions of interest. The connections between two regions are important for the characteristics of the brain network. Detecting such a subnetwork in the brain can help analyze which set of nodes is responsible for differences between groups with regards to a particular target variable.

All nodes in the input graph, $G(V, E)$ were given a constant negative weighting of $\{-0.01\}$ and all the edges obtained were positive since the absolute values of the coefficients in subsection 3.3.1 were taken. The small negative penalty was given to ensure a negative penalty on the inclusion of an extra node. Since all the edge weights were positive, the negative weights of the nodes ensured that the overall positive effect of the edge weights overcame the cost of keeping an extra node. This ensured the conservation of a specified number of nodes alongside compatibility for MEWIS (Eq. 3.3).

In the modification, giving the nodes a constant weight of zero results in all nodes getting preserved. This happened since there were no negative edges so none of the edges get eliminated. In this case the sum of all original edge weights remains the highest according to Eq. (3.3). This effect was observed since the preprocessing step of the implementation in Loboda et al., 2016 was disabled and all nodes in the input graph remained well connected. Similarly, assigning any type of positive node weight is incompatible for the implementation in subsection 3.4.2 since no nodes and edges would get eliminated.

The feature selection was carried out by making an input graph and a corresponding output graph. For each type of feature (mean FA, mean streamline length or number of streamlines), a graph was created specific to the target label of interest. The nodes in this graph correspond to ROIs mentioned in the appendix. The edges were the connections between the ROIs weighted by the statistical coefficients calculated for the training data with respect to the target variables. Such a graph was termed as the 'input graph'.

Sparsity to the input graph was introduced using two constraints. The absolute value of the edge weights (raw features represented by statistical coefficients, subsection 3.3.1) shall not be zero. Simultaneously, features are conserved if and only if the tractography of each subject contains at least one streamline between the two nodes.

For creating the output graph, parts of the pipeline used in Loboda et al., 2016 were modified. Constraints mentioned in section 2.4 were extended and the additional constraint was to preserve a specified number of nodes using the constraint:

$$\sum_{v=1}^{V} y_v = m \qquad \forall v \in V \tag{3.4}$$

where m is a controllable parameter specifying the number of nodes needed to be preserved in the output graph and y_v is the binary variable mentioned in Equation 2.22. The preprocessing module of the MCWS Java solver was disabled due to it's incompatibility with the required constraint (Eq. 3.4). This makes the subgraph produced an induced subgraph.

An object oriented approach was taken to implement a class derived from the *networkx* (Hagberg et al., 2008) for the different input graphs and their corresponding output graphs. At any time, the graphs created for a specific use case could be read from text files and their properties could be explored.

In order compare the features selected with this technique to the baseline experiments, it was important to infer the number of edges being preserved based on a specific input configuration. The output number of edges was then analyzed as function of the number of nodes. The results are presented in Fig. 4.6.

3.5 Supervised Classification

A Randomized Search 5-fold cross validation with implementation from *scikit learn* was carried out for different algorithms in different cases based on the type of feature, the feature selection technique, target variables along with target labels are presented in Figure 3.3. The cross validation was refitted using balanced accuracy on the test set. After the randomized search was finished, the best estimator was taken and fit to the training set. This classifier was then used to make predictions on an independent test set with difference in age range.

The training set consisted of 141 subjects, 83 females and 58 males aged 26-30. The independent test set consisted of 62 subjects, 34 males and 28 females aged 22-25. The different age ranges for the training and the test set were chosen in since the HCP data contains scans of twins, but this information was available only from the restricted access of the data. Grouping the test and training set according to age ensured that both the twins remain in either the training set or the test set. If it was not done this way then the classification performance would become higher only due to similarity in features of one twin to another.

One model was trained each time there was a different configuration of settings presented in Figure 3.3. Classification performance was assessed using the metrics AUC, f1-score, accuracy and balanced accuracy.

4 Results

In this section, the quantitative and visualization results from the methods presented in chapter 3 will be presented. The first two sections will present results from preprocessing of the DWI images mentioned in subsection 3.2.2 and extraction of connectivity matrices. The subsequent sections will describe the nature of the acquired data as well as classification results based on the comprehensive classification parameters illustrated in Figure 3.3.

4.1 Preprocessing Visualization

During the process of generating the connectome (subsection 3.2.3) visual inspection was required to investigate the properties of different types of images produced. The results illustrated in Fig. 4.1 were used to evaluate the adherence of the implementation to the conceptual framework mentioned section 3.2.

The 5TT image obtained after tissue segmentation (subsection 3.2.1) was 4D and could not be easily displayed in 3D. In Fig. 4.1.a this 4D image is displayed according to gray-scale mapping in 3D. The intensity of the background was set to 0, that of cortical gray matter at 0.5, subcortical gray matter at 0.75, white matter at 1, CSF at 0.15 and pathological tissue at 2. With this gray-scale mapping, it was visually observed that the 5TT images did not contain any erroneous labels.

Three different volumes containing the tissue densities of WM, GM and CSF are obtained by determining the corresponding fODFs using Multi-Shell Multi-Tissue Constrained Spherical Deconvolution (MSMT-CSD). The three images were then concatenated to a 4D image with RGB encoding: CSF as red, GM as green and WM as blue. In Fig. 4.1.b, the 4D image is visualized as an RGB. Using an interactive window from ***mrview*** (J.-D. Tournier, R. Smith, et al., 2019) the RGB image was zoomed and checked for compliance of the response functions with the fODFs and tissue segmentations.

After obtaining information about the tissue segmentations and fODFs, one million fibers whole brain tractography was generated (Fig. 4.1.c). In this tractogram the red fibers run in the right to left direction, green ones in anterior-posterior and blue in head-to-foot direction (Hobert et al., 2013). Finally, the connectome was

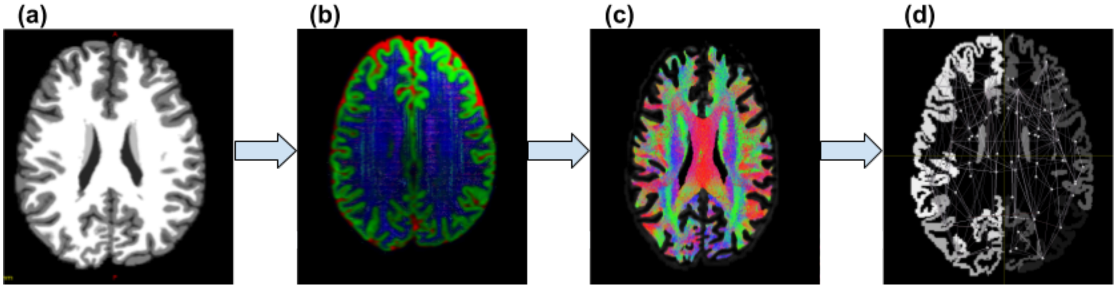


Figure 4.1: Visualization of pipeline used to create a connectome for each subject on an axial slice (a) Five tissue segmented image visualized in grayscale. (b) A slice of a 4D image mapped in 3D using RGB encoding tissue densities, CSF as red, GM as green and WM as blue. (c) ACT of one million fibers overlaid on an axial slice of the brain. (d) Visualization of the spatial location of the center of masses of different ROIs.

visualized (Fig. 4.1.d) with the nodes representing the center of mass of different gray matter regions and the edges as the connections between them. The intensity of this image is not equal to the signal intensity but the number assigned to the ROI, so any ROI that has a higher numbering in the lookup table is shown brighter (see Appendix).

Overall, visualizing the preprocessing pipeline helped to evaluate the anatomical correspondence of the tractography. It also helps understand the relationship between the generated connectome and white matter structural brain connectivity.

4.2 Connectome Visualization

After the connectivity matrices for all subjects were generated using the methods in section 3.2, a group averaged connectome was created. The raw connectivity matrix was dense, containing $n = 84$ nodes and $n_e = 3486$ edges. The edge weights were widely distributed for all the three connectivity metrics: mean FA, mean streamline length, and number of streamlines. Visualizing the connectome as a simple connectivity matrix or a heatmap was hence not an effective strategy.

A chord plot was chosen to represent structural brain connectivity because of the ease of representing gray matter nodes along the boundary of the graph. It is an effective way of visualizing the flow between different nodes. The number of streamlines could be easily encoded within a chord plot since the number of arcs drawn between any two nodes can be specified. The reason for the selection of the number of streamlines as a connectivity metric for structural brain connectivity representation will be presented in subsection 4.3.1.

Figure 4.2: A chord plot representing the group averaged connectome for all subjects. Nodes at the end of the circle correspond to ROIs. The thickness of each edge corresponds to the number of arcs between two ROIs. The number of arcs is proportional to the number of streamlines. In this interactive visualization the selected node (middle frontal gyrus in the left hemisphere) and its corresponding edges are highlighted while the other proprieties are made transparent. Edges connecting the selected node and target nodes are colored according to the target nodes. For example, the arcs are blue colored when there are connections to the precuneus in the left hemisphere.

The Fig. 4.2 has been created with help of package *holoviews* (Stevens et al., 2015). In this figure, average number of streamlines (for all subjects) between any two regions are visualized, $n = 65$ nodes are shown. Each node is assigned a different color according to the color palette in *holoviews* (Stevens et al., 2015). The arcs in the graph have been obtained by thresholding the group averaged number of streamlines between any two ROIs. This threshold was selected to be $N = 900$. Only the connections which have more than 900 streamlines are shown. The plot represents the number of streamlines along the preserved connections scaled down by a factor of 10 for ease of representation.

The detailed coloring scheme used for creating the chord plot gives added visual cues about the source and the target nodes. Once a node is selected the edges get colored according to the color of the target node. As an example, the cortical region caudal middle frontal in the left hemisphere is selected in Fig. 4.2. It has prominent connections to the cortical regions superior frontal, rostra middle frontal and precuneus in the left hemisphere. A weak connection to the cortical region parsopercularis is also well represented by this diagram.

Overall, this comprehensive diagram is a systematic representation of whole brain connectivity. It can be used to analyze brain connections at the subject level or the group level.

4.3 Feature Analysis

Once the connectivity matrices were ready, it was important to investigate an effective connectivity metric and determine importance of within ROI connections. From the preprocessing pipeline presented in subsection 3.2.3 there were three types of connectivity metrics obtained for each subject: number of streamlines, the mean streamline length and the mean FA. Each of these features encodes a different biological property of ROI-to-ROI connections and gave different classification results. There is no current consensus on which feature qualifies as an effective measure of connectivity between any two nodes (Yeh et al., 2020). The number of streamlines was selected to be the major focus of the classification tasks due its superiority in classification performance, which will be illustrated subsequently. Further, exclusion of self loops from the connectivity matrices will be experimentally justified.

4.3.1 Differences in Connectivity Metric

As mentioned above, it was important to explore different types of connectivity metrics obtained from the connectivity matrices constructed using the methodol-

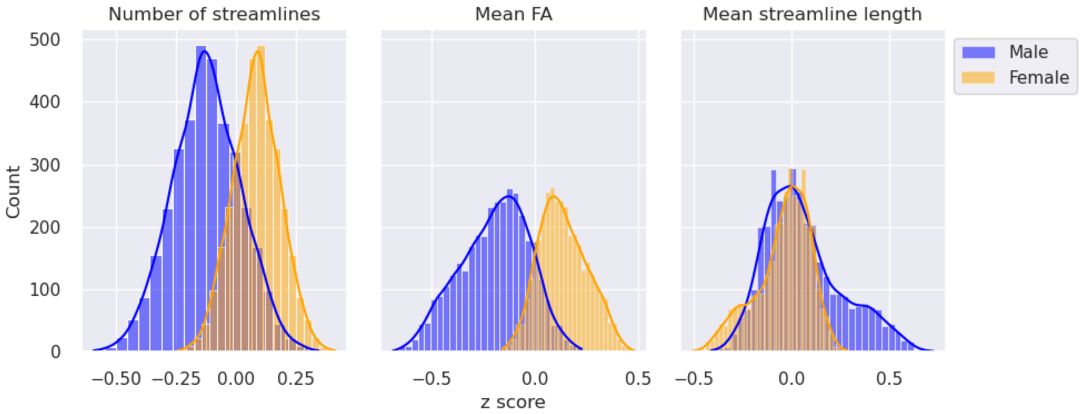


Figure 4.3: Three overlaid histograms representing the frequency of mean z-scores for different types of features. The orange histograms are for females and blue ones are for males. The y-axis represents the number of between ROI connections.

ogy illustrated in Fig. 3.1b. Evaluation of z-score distributions of ROI connections separately for males and females was taken as a method for data evaluation. Two overlaid histograms for each of the three different connectivity types helped speculate how well they separate between the two classes.

The histograms were formed using data (containing both data samples from males and females) in the training set. Each feature was standardized by removing the mean and scaling it to unit variance. After scaling the feature value for each subject could be represented by a z-statistic. As expected, the mean of the z-statistic (of one feature) was zero for all training subjects. However, the differences were prevalent when separate means were taken separately for the two genders. Hence, for each connection there was one z-score for males and another for females depending on the connectivity metric. For each connectivity metric, there are 3486 features considering pairwise connections excluding self loops (subsection 3.3.2) in an undirected graph. This number corresponds to the total count represented in Fig. 4.3.

From these distributions, it can be inferred that the number of streamlines have characteristics that can help interpret gender differences in structural brain connectivity. Its histograms show a smooth Gaussian distribution for males and females, and has significantly higher peaks than those for the other two features. Even though the mean FA features have a smaller overlap visually, the spread of its histograms is greater than that for the streamline count. The histograms of the mean streamline length are overlapping and skewed.

From the differences in z-score distribution, it can be inferred that women on an average have more number of streamlines as compared men in view of the same connections. This observation is coherent with evidence in Szalkai et al., 2015 showing that women have more densely connected brains than men. Further, women have higher white matter to gray matter ratio and hence the number of streamlines detected for women is higher (Taki et al., 2011).

The mean FA plots illustrate that women have higher mean FA values than men. An inconclusive trend is seen in the case of mean streamline length. Currently, there are no conclusive results for benchmarking such a distribution of FA values, most results depend on the ROIs being considered (Kanaan et al., 2012). It is difficult to conclude that female brains have higher FA values since existing studies in whole brain connectivity do not provide conclusive results (Ingalhalikar et al., 2014). The mean streamline length remains a relatively unexplored connectivity metric hence substantial reasoning cannot be provided for explaining the trend of streamline lengths in 4.3 (Bajada et al., 2019). The difficulty of interpreting the mean FA as well as mean streamline length is detailed in section 5.1.

The results for the number of streamlines are coherent with *a priori* knowledge and evidence in literature. The tractography was carried out in individual subject's space which could have caused biases in the comparison of fiber count between different subjects. Such biases were corrected for comparing the streamline count or length for different subjects using informed filtering from the SIFT algorithm (Yeh et al., 2020). The SIFT algorithm maps the reconstructed fODFs back to the original dMRI data. It then modifies the streamlines in accordance with image data and hence improves the biological correspondence of the streamline count. With this model, the number of streamlines becomes proportional to the total cross-sectional area of the white matter fibers connecting any two ROIs. This proportionality makes the streamline count an interpretable connectivity metric as the the cross-sectional area of white matter pathways is of high anatomical correspondence (R. E. Smith et al., 2015a).

The experiments stated above and classification results in subsection 3.4.1 help establish the advantage of streamline count as a connectivity metric as compared the other two metrics. It gives a biologically coherent measure of brain connectivity and achieves better performance than mean FA and mean streamline length.

4.3.2 Self Loops

Evaluation of the importance of self loops was carried out on the basis of the experiment mentioned in subsection 3.3.2. The results in Table 4.1 were obtained by training the baseline experiments with and without the exclusion of self loops. The report values correspond to classification performance on the test data. A positive t-statistic in this case indicates that the classification performance is greater with

Label	**Metric**	**Feature**	**T statstic**	**P value**
Personality Traits	Balanced accuracy	Mean FA	-0.196	0.846
		Mean streamline length	0.156	0.876
		Number of streamlines	-1.248	0.217
	Area under ROC curve	Mean FA	0.071	0.944
		Mean streamline length	1.332	0.188
		Number of streamlines	-1.42	0.161
Gender	Balanced accuracy	Mean FA	0.396	0.695
		Mean streamline length	2.068	0.048
		Number of streamlines	0.98	0.335
	Area under ROC curve	Mean FA	0.491	0.627
		Mean streamline length	3.875	0.001
		Number of streamlines	2.199	0.036

Table 4.1: Results for a paired samples t-test on the test set. The paired samples are the classification metrics of the data with and without the inclusion of self loops. The p values correspond to the t-test of two set of observations for the same classification metric constructed using methodology in subsection 3.3.2.

the inclusion of self loops. Meanwhile, a negative t-statistic indicates the opposite trend. Statistical inference on this data indicates that the exclusion of self loops from the data analysis does not have a statistically significant effect on classification accuracy. The statistics obtained from the analysis of both the target labels, i.e. the personality traits and gender labels do not lie in a significant confidence interval. The p-values in Table 4.1 are indicative of these results with most of them $p > 0.05$.

Consider the statistics for the mean FA feature. All the p-values are $p > 0.05$ and the t-statistic has opposite signs when it comes to considering two cases for personality classification. The first being balanced accuracy and the second as the AUC. The t-statistic for the first case is $t = -0.196$ and for the second case is $t = 0.071$. This opposite trend for two different measures is indicative of the fact that there is no general trend of classification performance.

With this investigation it became clear that the self loops can be eliminated from the data of all subjects. This ensured that the MEWIS solver based technique as well as the baseline filters features from the same set of raw data.

4.4 Baseline Analysis

The baseline analysis mentioned in subsection 3.4.1 with the parameter k denoting the percentage of features to be preserved chosen to select the best performing connectivity metric. Initally, $k \in \{2, 5, 10, 50, 100\}$ was chosen to see an overall trend of classification performance. One important research question to be explored with this analysis was to determine whether doing any feature selection gives an added advantage for classification performance.

From the results in Fig. 4.4 it became evident that the number of streamlines served as a good metric for gender classification. There was only a loss of around 5% performance for AUC while reducing the percentage of features from 100% to 5% for both types of edge representations i.e. the t-test and f-scores and all three classifiers. In terms of number of features this meant that choosing just 175 out of 3486 features most of the discriminative information could be preserved.

Another consideration for gender classification was that even though the p-values of the t-test were scaled according to the function $f(x) = (-1) \times log_{10} p_x$ and the raw f-scores were used for scoring, the performance of the baseline analysis remained similar. In our training data the number of males and females were not equal and hence a Welsch's t-test assuming similar variances of the feature for the two groups was applied (subsection 3.3.1). This could have introduced slightly different ranking from the f-scores as compared to the t-test.

The classification of the different personality traits was more complex as compared to gender classification. All the possibilities of connectivity metrics and five different personality traits with Pearson correlation coefficient as the feature scoring were considered in the pipeline. However, an overall trend could be observed only with neuroticism and extraversion. For the traits openness and conscientiousness, the classification performance was often worse with feature selection than without. Hence, it could not be concluded that doing feature selection for personality traits is profitable. Only a subset of the experiments with the neuroticism and extraversion personality traits using streamline count as valid connectivity are mentioned according to subsection 4.3.1. From the results in our analysis, it could be concluded that personality traits cannot be accurately classified using structural brain connectivity.

In Fig. 4.5 for neuroticism, there is a trend of increase in performance with only a subset of original features using all three classifiers. The trend for the MLP shows greater than 65% AUC for top 5% of features while 45% without any feature selection. Classification performance on extraversion label increases with decrease in the number of features for classification with MLP and SVC. The strongest trend seen with the SVC, the AUC increases from 50% (without feature

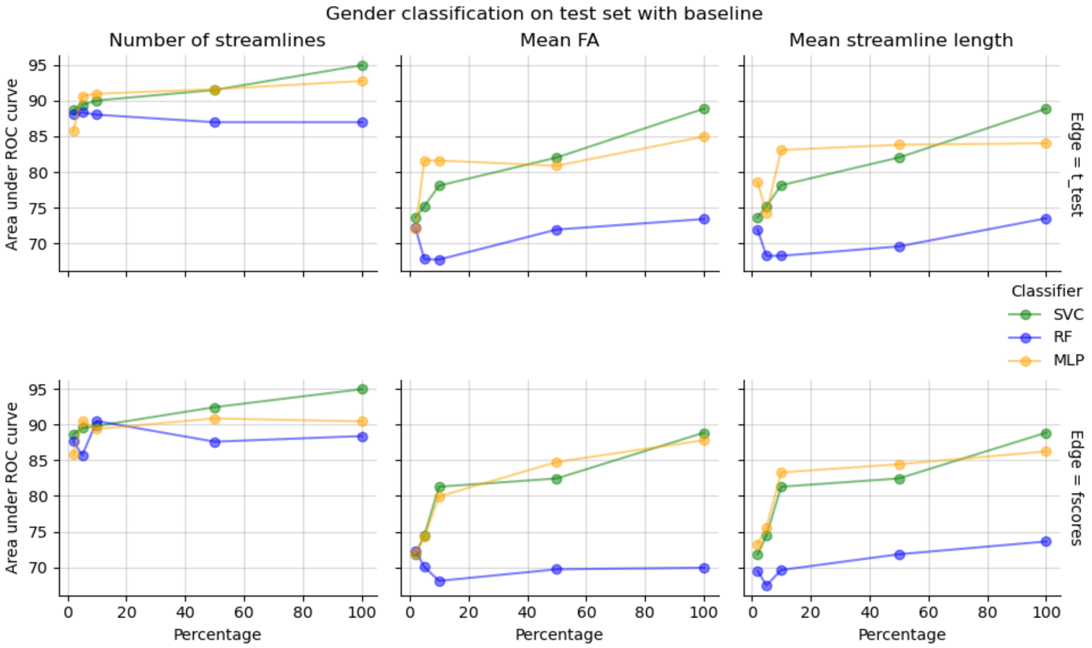


Figure 4.4: Baseline analysis for gender classification. The AUC is represented as a function of percentage of features preserved. The three columns correspond to three different metrics used to represent structural brain connectivity. For the experiments illustrated in the first row the edges are filtered according to the p-value of the t-test and in the second row filtering is done using f-scores.

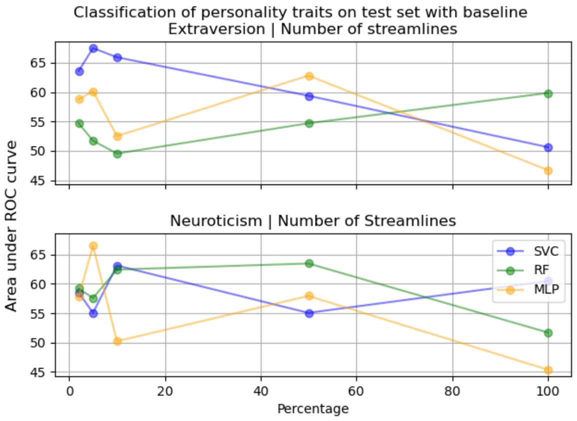


Figure 4.5: Baseline analysis on the basis of personality traits using Pearson correlation coefficient as a filter method. For neuroticism, there is a trend of increase in performance with only a subset of original features using all three classifiers. For this particular configuration feature selection lead to much improved classifier performance with all three classifiers. A steep increase can be seen for extraversion where the AUC increases from about 50% (random chance) to about 68%.

selection) to around 67% (using top 5% of the original features). The overall trend for neuroticism and extraversion illustrates that it is useful to do some feature selection in order to predict these traits

For gender classification, an interpretability-performance trade-off was observed. This trade-off is between the interpretability inversely proportional to the number of features and classifier performance that is directly proportional to the number of features. Feature selection seems profitable since the interpretability increases many fold when the number of features reduces from 3486 to just 175.

The case for classification of personality traits is however different. Feature selection seems profitable in some cases but not all. In this section only the configurations for which feature selection shows an added advantage have been shown. For the selected configurations, classification based on the features selected by the MEWIS solver features is significantly better than the non-informed baseline feature selection.

4.5 Tracing Back to Original Features

A major motivation behind implementing the MEWIS based technique for this thesis was interpretability of the feature selection. In this section the interpretability of the set of features selected by the baseline as well as the solver based implementation are compared. It is observed that the solver based method is superior to the traditional baseline method due to connectedness of the output features.

The two subsequent subsections present visualizations for the features selected by both the feature selection techniques. Furthermore, the subsection 4.5.1 provides an insight into how the solver implementation preserves edges.

4.5.1 Subgraph Features

For classification with the solver based technique detailed in section 2.4, the number of nodes preserved could be specified using a parameter m. Determining the order of growth in the number of edges as a function of m was important to trace back the MEWIS selected features to the original *dataframe*. Figure 4.6 shows a quadratic form with the number of edges y as a function of the number of nodes x. This behavior is expected as the graph is supposed to be connected; i.e. there exists an edge between any two nodes in the graph. Using these results, the extent of interpretability from the solver results could be controlled.

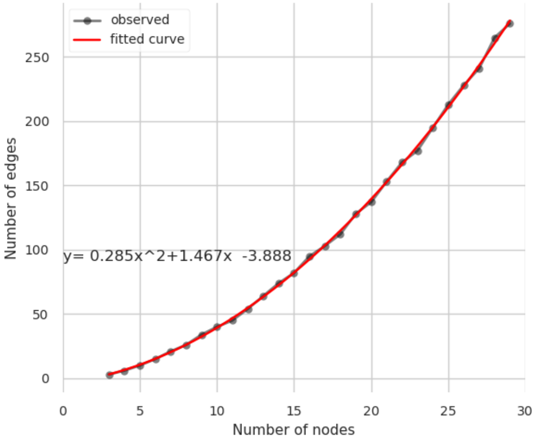


Figure 4.6: The number of edges preserved by the MEWIS solver as a function of the number of nodes supplied to the parameter m (Eq. 3.4). The black dots illustrate the number of edges preserved for inputs $m \in [4, 30]$. A quadratic curve was fitted to map this trend. The red line represents the fitted model with the parameters mentioned in the figure. Using the fitted polynomial the number of edges preserved could be predicted before running the solver. This saved computational time and effort since the nature of output graphs could be predicted beforehand.

4.5.2 Baseline-Solver Comparison

In order to compare the two feature selection techniques, a separate input graph had to be created for a two use cases according to Figure 3.3. These two use cases were visualized to compare the interpretability of predictions from the solver and baseline analysis.

In the first use case, the type of feature is number of streamlines, the target label is gender, edge representation is f-score and feature selection is solver with 10 nodes preserved in the subgraph. In the second use case, all the parameters were the same as the first use case except for baseline as the feature selection technique and number of edges as 41 for comparison with 41 edges preserved in the first use case (using polynomial in Fig. 4.6). The visualizations of these two use cases are Fig. 4.7 and Fig. 4.8. In both the visualizations, the edge widths represent the f-scores and the color represents the group averaged number of streamlines. The group averaged number of streamlines and f-scores are determined on the training data only.

In Fig. 4.7 the MEWIS solver based implementation reduces the input graph and preserves 10 nodes along with 41 edges. It can be seen that the graph is well connected and structured. 6 out of 10 most important nodes are subcortical gray

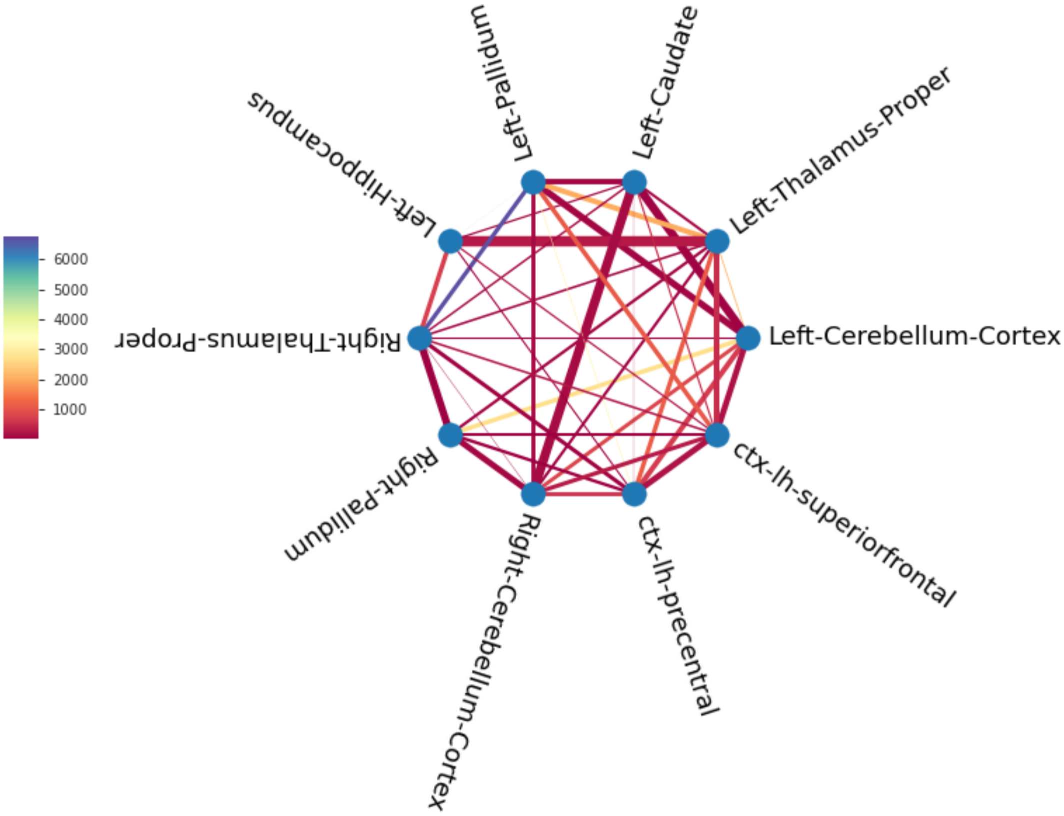


Figure 4.7: 10 most important nodes out of out of 84 determined using the MEWIS solver for gender classification. The edge widths represent the f-scores of and color represents the number of streamlines between the two regions. A higher number of streamlines does not directly correspond to a higher f-score. 6 of the 10 selected nodes are subcortical regions; the thalai, the pallidai, the left hippocampus and left caudate while the remaining are cortical regions. This figure depicts that gender differences might be more prominent in subcortical brain networks than cortical brain networks.

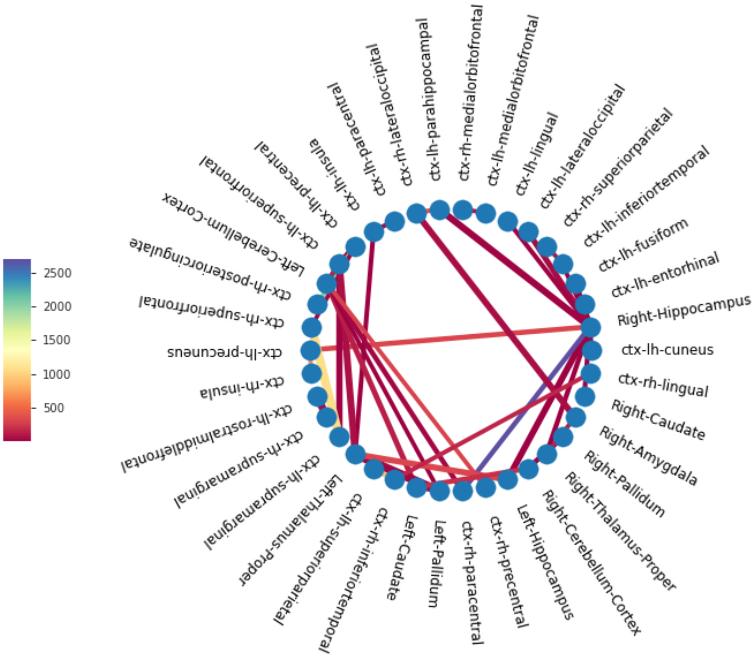


Figure 4.8: 41 most important features, and consequently 38 nodes preserved using baseline analysis for gender classification. The edge widths represent the f-scores and color represents the number of streamlines between the two regions. This figure shows that the baseline analysis does take graph topology into account and the features selected by this technique are not informative about any patterns such as the ones inferred using MEWIS solver in Fig. 4.7.

metter regions. These 6 regions are the thalai, the pallidai, the left hippocampus and left caudate. The lookup table from the Desikan-Killiany Atlas had total 68 nodes for cortical gray matter regions and 16 nodes for subcortical gray matter (see Appendix). The ratio of selected nodes for subcortical regions 0.375 is higher than for cortical regions 0.058. Such a pattern indicates that gender differences might be more prominent in networks of subcortical regions rather than cortical ones. Figure 4.8 exemplifies the fact that the baseline method selects the same number of edges i.e. 41 as the solver by preserving 38 nodes, almost four times the nodes preserved by the solver. The graph is not connected and contains isolated edges, very often nodes with a degree of only 1 are seen.

Figure 4.7 and Fig. 4.8 illustrate the superiority of the MEWIS solver generated subgraphs over baseline features in terms of interpretability. In a research setting, the MEWIS solver method would be more beneficial since it provides more interpretable results. With this technique it becomes easier to determine the subnetworks related to a particular target variable.

4.6 Model Performance

Classification performance on the target label varied depending on the parameters of the data given to the classifier according to Figure 3.3. This section will present the classification results for solver and personality traits respectively. For the classification of gender, the comparison of classification performance was easily interpretable. At first, the baseline experiments for $k \in \{2, 5, 10\}$ were compared with solver nodes $m \in \{5, 10, 15, 25, 30\}$ due to computational complexity of the exhaustive search. Figure 4.9 was used to determine the which classifier performs the best and gives results free from overfitting.

The SVC based classification produced the most stable results out of the three classifiers used and was taken as the choice for an exhaustive search. Once the classifier was decided upon, classification performance for a higher number of data points was determined. This time the parameter top k percentile of features to be preserved with the baseline was decided by using the polynomial in Fig. 4.6 so that the number of edges selected by the solver and the baseline are similar and classification results can be easily compared.

The subsequent results using SVC (Fig. 4.10) clearly indicate a cutoff number of edges, i.e. precisely 137 (obtained from the function in Fig. 4.6). The MEWIS solver does better than the baseline analysis in terms of classification performance after the cutoff value of 137 edges. Even for the number of edges below 137, the classification performance of both the methods remains comparable. It can be

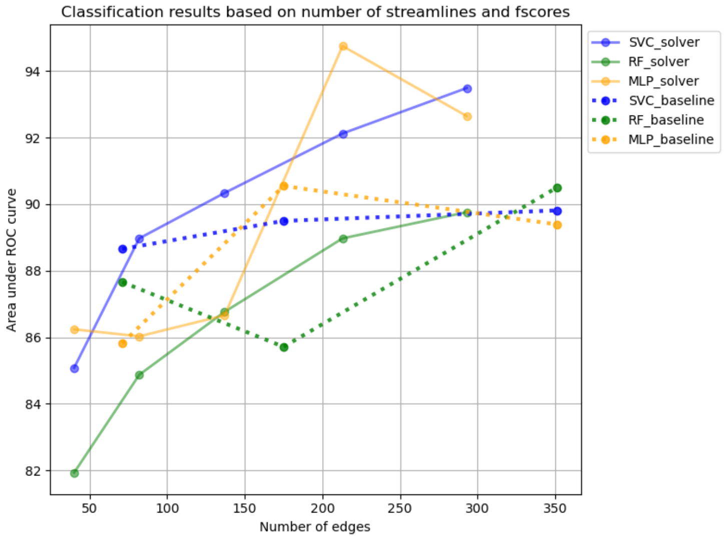


Figure 4.9: Comparison of AUC for gender classification on the independent test set. The performance of three classifiers SVC, RF and MLP on basis is compared. The solid lines and dashed lines represent the performance for features filtered according to the solver and baseline experiments respectively.

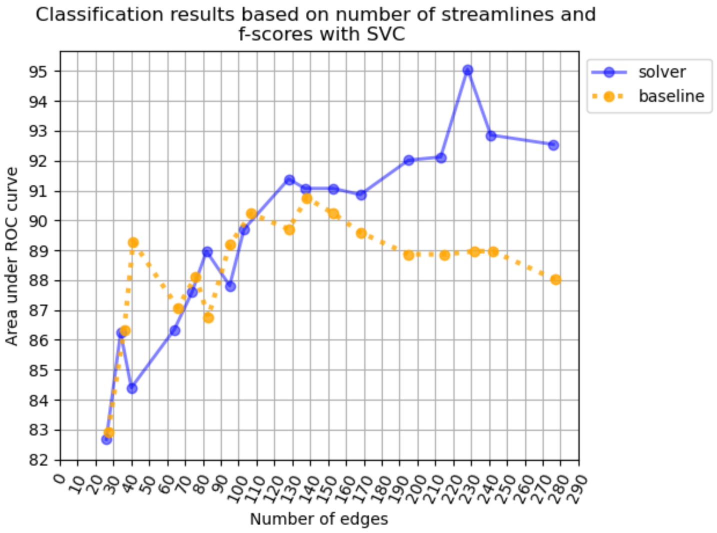


Figure 4.10: Comparison of solver and baseline for gender classification with almost the same number of features selected in each case. The streamline count is used as the connectivity metric and f-scores are used as the feature scoring technique.

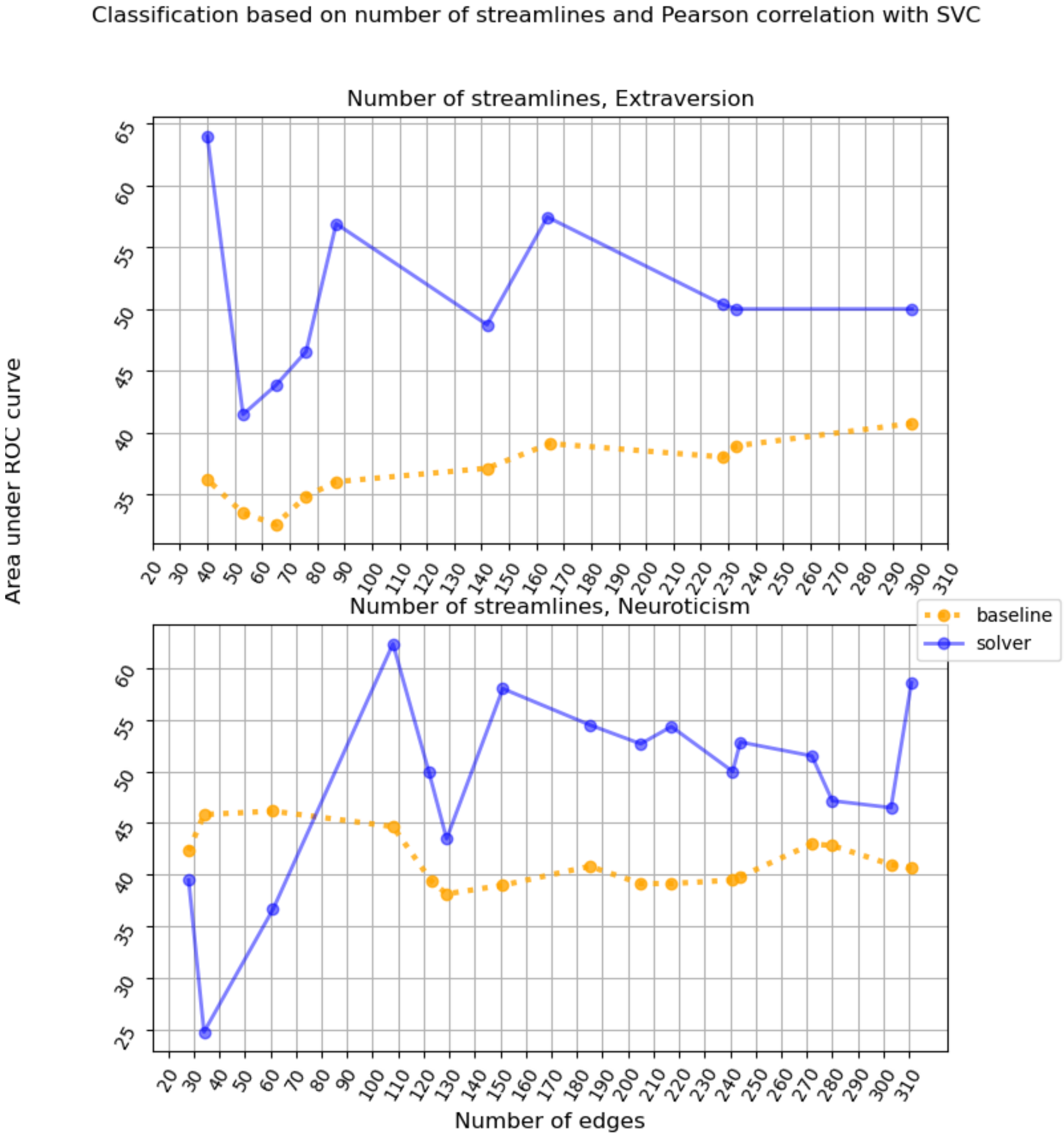


Figure 4.11: Comparison of solver and baseline for classification of personality traits using SVC. This figure represents an exhaustive search carried out for only two out of five personality traits with the streamline count as the connectivity metric and Pearson correlation as the feature scoring technique. The cases represented in this figure have been selected due to the reasons mentioned in section 4.4.

Hyperparameter	Solver	baseline
C	5.9	0.522
Gamma	4e-4	8e-4
kernel	rbf	rbf
class weight	balanced	None

Table 4.2: Cross validation parameters for SVC trained for gender classification. The parameters which give the best area under the curve for the use case mentioned in the section above have been presented

inferred that preserving more than 137 edges or 20 nodes (according to Fig. 4.6) with the MEWIS offers more interpretability as well as better classification performance than the baseline method. This means that with $m > 20$ nodes, there is a decrease in interpretability but an increase in classifier performance. Preserving less than 20 nodes leads to a reduction in classifier performance but an increase in ease of visualization and interpretability.

The use cases in Fig. 4.5 was used to carry out an exhaustive search for the comparison of model performance between the two feature selection techniques. Fig. 4.11 highlights this comparison. The MEWIS method performed much better than baseline analysis throughout the range of feature numbers. For the classification of neuroticism the solver is able to achieve a maximum AUC of around 64% which is a better than the maximum score of 40% by a margin of 24%. Similarly, for neuroticism the SVC achieves about 62% accuracy with the MEWIS solver and a little over 45% for the baseline. One important detail to note is that the baseline performance might not exactly be the same as that in Fig. 4.5 since the number of features used for classification is not the same.

On the basis of the classification results for personality traits and gender, it can be concluded that the MEWIS solver is a better feature selection technique than the baseline analysis.

In the remainder of this section, the parameters of the best estimators for gender classification will be presented. The SVC gave most stable results as compared to other classifiers. The best performing estimators are determined using Fig. 4.10. For gender classification using the solver, the maximum AUC of 95% for the independent test set could be achieved. This performance is obtained for 237 edges corresponding to 26 nodes. Using the baseline technique, the best estimator could achieve 91% area under curve for around 138 edges.

In Table 4.2 the parameters of the best estimators for the MEWIS solver based method and baseline used for gender classification using number of streamlines and f-scores will be presented.

5 Discussion

In this work classification based on a subgraph derived from structural brain connectivity was the major challenge. As mentioned in Figure 3.3, a large combination of possibilities for designing the classification task was investigated to obtain the most interpretable and statistically significant results. The goal was to not only gain the maximum classification performance but also to design a generalized framework in which each step can be visualized. The architecture remains specific to structural brain connectivity but is generalizable enough to predict different target variables. Major considerations for the task design such as connectivity features and graph structure are explained in the subsequent sections.

5.1 Connectivity Metric

This section will discuss in detail the reason for directing the analysis in chapter 4 towards using the streamline count as a connectivity metric. It will also elucidate the inefficacy of other connectivity features in terms of biological correspondence.

There were three types of features used as representatives of structural brain connectivity. The number of streamlines, the mean FA of streamlines, and the mean streamline length. Generally, the number of streamlines seemed to have performed much better than the other two features considering the AUC for classification of gender on an independent test set (subsection 3.4.1). It performed better with both baseline and solver based methods, regardless of the labels i.e. personality trait or gender.

The superior performance of the number of streamlines using the MEWIS solver based approach seems plausible due to the incorporation of a particular constraint in the pipeline. A feature from the connectome is included in the input graph for the solver if and only if there is at least one streamline present for all subjects. Such a condition remained unique to the number of streamlines. Perhaps trying such conceptual *a priori* knowledge with the mean FA and mean streamline length could yield different results. Furthermore, the streamline count was a biologically coherent measure of brain connectivity as explained in 3.4.1.

The mean FA is a local metric derived from one voxel at a time. Interpolating such a local measure to infer connectivity between brain regions is hence problematic. The mean FA is modulated by intra-voxel orientational dispersion, myelination, packaging density, etc. A linear relationship between structural brain connectivity and mean FA cannot be established since a change in anisotropy does not necessarily correspond to a change in FA (D. K. Jones et al., 2013).

In our analysis, the number of fibers was downsampled from five million to one million by using the SIFT algorithm to generate anatomically constrained tractography. This might have caused shorter fibers to be over-represented (R. E. Smith et al., 2015b). The non-uniform distribution of the streamline lengths is presented in Fig. 4.3. Other biases in the tractography such as streamline to node assignment make the streamline length an abstract measure connectivity (Yeh et al., 2020).

5.2 Constraining the Number of Nodes

As mentioned in subsection 3.4.2, the MEWIS solver implementation was give a constraint to preserve a specified number of nodes in order to indirectly control the number of edges preserved. Without this constraint, an induced subgraph could not be formed and the results produced by the solver could not be compared to the baseline exhaustively. Forming an induced subgraph is of high importance in clinical applications in which brain networks relevant to a disease pathology have to be found. Translating such subgraph based classification from computational models to research practice requires the need to compare the performance with existing methods.

A subgraph $\tilde{G}(\tilde{V}, \tilde{E})$ of $G(V, E)$ is called an induced subgraph if and only if two vertices u, v adjacent in $\tilde{G}$ are also adjacent in $G(V, E)$. In the implementation of MEWIS for this project, induced subgraph is formed by deleting vertices (with the maximal sum of edge weights) but alongside ensuring that the subgraph is connected.

Without constraining the number of nodes or forming an induced subgraph, the output graph returned from the solver had an arbitrary number of edges based on the original implementation in Loboda et al., 2016. When the parameter m for specifying the number of nodes was added different output graphs could be generated from the same input graph. These different output graph had a different number of edges on the basis of the number of nodes. All these graphs contained a different number of edges which could be matched to the k percentile of features preserved by the baseline. Hence, the incorporation of the constraint on the number of nodes was pivotal to evaluate the performance of the features selected by the MEWIS solver.

5.3 Personality Classification

The classification of personality traits was a methodologically challenging task. The personality traits are continuous in nature. A regression task was not formulated for these continuous variables since the regression task would be incompatible with the pipeline. The major focus of this thesis was to formulate a discriminative subgraph.

There are no population parameters that quantify the mean of a personality trait or its standard deviation. Personality traits are inferred mostly on the basis of questionnaires which introduces an abstraction in their use as a biomarker. It was difficult to predict personality since it depends on various factors which might not be elucidated by structural brain connectivity. Most studies using structural MRI data for the study of personality have failed in terms of replicability (Dubois et al., 2018).

According to our methodology, binning the continuous personality traits for a classification task lead to an information loss. An attempt was made to incorporate such a loss by using the Pearson correlation coefficient that captures linear relationships. The size of our training set of 141 subjects might have limited the ability of the pearson correlation coefficient to model a relationship between the features and a particular personality trait.

Theoretically, the Big 5 personality traits are supposed to be orthogonal and hence they were classified one trait at a time in our pipeline. However, the personality traits have been found to be correlated to one another (Blackburn et al., 2004). The classification accuracy with the MEWIS might be different if a comprehensive model of personality is taken into account.

5.4 Triviality of Graph Kernels

Graph kernels are kernels defined on graphs. They are often used for classification from networks structured data such as molecules. However, they are not suitable for classification from structural brain connectivity and were excluded from the analysis pipeline. This remainder of this section describes the nature of graph kernels and why they are ill-suited for the task described in this thesis.

Graph kernels measure the topological similarity between different networks of the same type. There are different types of graph kernels which are widely used in bioinformatics and chemoinformatics. For example subgraph matching kernels are widely used in chemoinformatics to identify similar molecules and the prediction of protein function (Kriege and Mutzel, 2012). However, they are not free of shortcomings. A limitation of graph kernels is their high computational complexity which makes them applicable only for small graphs. Most graph kernels are built on unweighted graphs and often rely on the assumption that all nodes in the graph are uniform.

For classification from structural brain connectivity it is required that specific characteristics of each node be taken into account, and to not lose out on information about the importance of connections (Jie et al., 2018). In Figure 3.1b the connectivity matrix consisted of 3486 features with each node labelled according to a lookup table (see Appendix). Using a graph kernel to determine pairwise similarities between subjects with the high number of features was computationally infeasible. In light of these reasons, graph kernels were not used for classification. The MEWIS solver method proved to be well suited to the problem of feature selection in connectomics.

6 Conclusion

Analyzing the properties of structural brain networks is a challenging task. Especially classification based on dense graphs derived from brain connectivity requires a well-assessed scientific methodology. The pipeline implemented in this work is tailored towards accurate classification from structural brain connectivity. Three major characteristics of the pipeline make it apt for usage in clinical as well as behavioral research. These aspects are the consideration of graph topology, visualization at each step, and a generalized approach.

First, the graph topology incorporated using the MEWIS technique provides interpretable and visualizable results. This framework enables effective classifier training and statistical analysis. The structure of the output subgraph can be controlled by specifying a number of parameters: the number of nodes to be preserved, the type of ranking, and the type of connectivity measure. Extracting a k-induced subgraph gives the advantage of preserving edge information along with node information over the GMWCS. Furthermore, the number of edges can also be controlled indirectly by specifying the number of nodes. This makes it a better alternative to the state-of-the art filter based methods used in neuroimaging studies.

Second, all the steps right from the preprocessing of the connectome to the features selected for model training have a visual analog. Detailed visualizations that help evaluate biological correspondence of the data preparation and selection have been incorporated into the pipeline. These visualizations aid the evaluation of significant biases that can influence the exactness of an *in-silico* model of brain connectivity.

Third, the pipeline is a generalized method for classification using structural brain connectivity. It can be implemented for continuous or categorical variables. Special consideration has been given to remove the effects of information loss for continuous variables. It is also classifier independent and computationally efficient. An extensive analysis can be carried out by combining the MEWIS feature selection technique with a wide range of classifiers.

Overall, the method is novel but is not comprehensive on addressing all the challenges that come in the way of modeling structural brain connectivity. It gives a controllable trade-off between model performance and interpretability. The pipeline could further be improved by encoding more *a priori* knowledge from neurobiology and consideration of a larger dataset and using different imaging modalities. Another important observation could be to see if the features selected by the diffusion matrices correspond to the features selection by using functional connectivity matrices formed using fMRI data.

References

[1] C. Backes, A. Rurainski, G. W. Klau, O. Müller, D. Stöckel, A. Gerasch, J. Küntzer, D. Maisel, N. Ludwig, M. Hein, et al. "An integer linear programming approach for finding deregulated subgraphs in regulatory networks". In: *Nucleic acids research* 40.6 (2012), e43–e43 (cit. on p. 20).

[2] C. J. Bajada, J. Schreiber, and S. Caspers. "Fiber length profiling: A novel approach to structural brain organization". In: *NeuroImage* 186 (2019), pp. 164–173. ISSN: 1053-8119. URL: `http://www.sciencedirect.com/science/article/pii/S1053811918320512` (cit. on p. 50).

[3] D. S. Bassett and O. Sporns. "Network neuroscience". In: *Nature neuroscience* 20.3 (2017), pp. 353–364 (cit. on p. 16).

[4] T. E. Behrens, H. Johansen-Berg, M. Woolrich, S. Smith, C. Wheeler-Kingshott, P. Boulby, G. Barker, E. Sillery, K. Sheehan, O. Ciccarelli, et al. "Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging". In: *Nature neuroscience* 6.7 (2003), pp. 750–757 (cit. on pp. 16, 21).

[5] Y. Bengio and O. Delalleau. "On the Expressive Power of Deep Architectures". In: *Algorithmic Learning Theory*. Ed. by J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 18–36. ISBN: 978-3-642-24412-4 (cit. on p. 27).

[6] R. Blackburn, S. J. Renwick, J. P. Donnelly, and C. Logan. "Big five or big two? Superordinate factors in the NEO five factor inventory and the antisocial personality questionnaire". In: *Personality and Individual Differences* 37.5 (2004), pp. 957–970 (cit. on p. 65).

[7] F. Bloch. "Nuclear Induction". In: *Phys. Rev.* 70 (7-8 Oct. 1946), pp. 460–474. URL: `https://link.aps.org/doi/10.1103/PhysRev.70.460` (cit. on p. 6).

[8] F. Bloch. "Nuclear induction". In: *Physical review* 70.7-8 (1946), p. 460 (cit. on p. 11).

[9] D. Brogioli and A. Vailati. "Diffusive mass transfer by nonequilibrium fluctuations:Fick's law revisited". In: *Physical Review E* 63.1 (Dec. 2000). ISSN: 1095-3787. URL: `http://dx.doi.org/10.1103/PhysRevE.63.012105` (cit. on p. 9).

[10] C. Burges and C. J. Burges. “A Tutorial on Support Vector Machines for Pattern Recognition”. In: *Data Mining and Knowledge Discovery* 2 (Jan. 1998), pp. 121–167. URL: `https://www.microsoft.com/en-us/research/publication/a-tutorial-on-support-vector-machines-for-pattern-recognition/` (cit. on p. 23).

[11] R. Casanova, C. Whitlow, B. Wagner, M. Espeland, and J. Maldjian. “Combining graph and machine learning methods to analyze differences in functional connectivity across sex”. In: *The open neuroimaging journal* 6 (2012), p. 1 (cit. on p. 23).

[12] K. A. Clark, K. H. Nuechterlein, R. F. Asarnow, L. S. Hamilton, O. R. Phillips, N. S. Hageman, R. P. Woods, J. R. Alger, A. W. Toga, and K. L. Narr. “Mean diffusivity and fractional anisotropy as indicators of disease and genetic liability to schizophrenia”. In: *Journal of psychiatric research* 45.7 (2011), pp. 980–988 (cit. on p. 12).

[13] B. A. Cociu, S. Das, L. Billeci, W. Jamal, K. Maharatna, S. Calderoni, A. Narzisi, and F. Muratori. “Multimodal functional and structural brain connectivity analysis in autism: A preliminary integrated approach with EEG, fMRI, and DTI”. In: *IEEE Transactions on Cognitive and Developmental Systems* 10.2 (2017), pp. 213–226 (cit. on p. 2).

[14] D. Colquhoun. “The reproducibility of research and the misinterpretation of p-values”. In: *Royal society open science* 4.12 (2017), p. 171085 (cit. on p. 19).

[15] S. Cook, T. Jarrell, C. Brittin, Y. Wang, A. Bloniarz, M. Yakovlev, K. Nguyen, L. Tang, E. Bayer, J. Duerr, H. Bülow, O. Hobert, D. Hall, and S. Emmons. “Whole-animal connectomes of both Caenorhabditis elegans sexes”. In: *Nature* 571 (July 2019), p. 63 (cit. on p. 17).

[16] P. T. Costa Jr and R. R. McCrae. “The five-factor model of personality and its relevance to personality disorders”. In: *Journal of personality disorders* 6.4 (1992), pp. 343–359 (cit. on p. 32).

[17] F. Dell’Acqua and J.-D. Tournier. “Modelling white matter with spherical deconvolution: How and why?” In: *NMR in Biomedicine* 32.4 (2019), e3945 (cit. on p. 14).

[18] M. Descoteaux, R. Deriche, T. R. Knosche, and A. Anwander. “Deterministic and probabilistic tractography based on complex fibre orientation distributions”. In: *IEEE transactions on medical imaging* 28.2 (2008), pp. 269–286 (cit. on p. 15).

[19] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, et al. “An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest”. In: *Neuroimage* 31.3 (2006), pp. 968–980 (cit. on p. 31).

[20] O. Dietrich, A. Biffar, A. Baur-Melnyk, and M. Reiser. “Technical aspects of MR diffusion imaging of the body”. In: *European Journal of Radiology* 76 (Mar. 2010), pp. 314–22 (cit. on p. 10).

[21] J. Dubois, P. Galdi, Y. Han, L. K. Paul, and R. Adolphs. “Resting-state functional brain connectivity best predicts the personality dimension of openness to experience”. In: *Personality neuroscience* 1 (2018) (cit. on p. 65).

[22] A. Fornito, A. Zalesky, and M. Breakspear. “The connectomics of brain disorders”. In: *Nature Reviews Neuroscience* 16.3 (2015), pp. 159–172 (cit. on p. 17).

[23] S. Gerhard, A. Daducci, A. Lemkaddem, R. Meuli, J.-P. Thiran, and P. Hagmann. “The connectome viewer toolkit: an open source framework to manage, analyze, and visualize connectomes”. In: *Frontiers in neuroinformatics* 5 (2011), p. 3 (cit. on p. 33).

[24] A. Ghosh and R. Deriche. “A survey of current trends in diffusion MRI for structural brain connectivity”. In: *Journal of neural engineering* 13.1 (2015), p. 011001 (cit. on p. 9).

[25] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen, and M. Jenkinson. “The minimal preprocessing pipelines for the Human Connectome Project”. In: *NeuroImage* 80 (2013). Mapping the Connectome, pp. 105–124. ISSN: 1053-8119. URL: `http://www.sciencedirect.com/science/article/pii/S1053811913005053` (cit. on pp. 29, 34).

[26] A. Hagberg, P. Swart, and D. S Chult. *Exploring network structure, dynamics, and function using NetworkX*. Tech. rep. Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008 (cit. on p. 42).

[27] M. Haouari, N. Maculan, and M. Mrad. “Enhanced compact models for the connected subgraph problem and for the shortest path problem in digraphs with negative cycles”. In: *Computers & operations research* 40.10 (2013), pp. 2485–2492 (cit. on p. 22).

[28] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009 (cit. on p. 25).

[29] HCP. “WU-Minn HCP 900 Subjects Data Release: Reference Manual”. In: (2015) (cit. on p. 29).

[30] M. K. Hobert, V. M. Stein, P. Dziallas, D. C. Ludwig, and A. Tipold. “Evaluation of normal appearing spinal cord by diffusion tensor imaging, fiber tracking, fractional anisotropy, and apparent diffusion coefficient measurement in 13 dogs”. In: *Acta Veterinaria Scandinavica* 55.1 (2013), p. 36 (cit. on p. 45).

[31] M. Ingalhalikar, A. Smith, D. Parker, T. D. Satterthwaite, M. A. Elliott, K. Ruparel, H. Hakonarson, R. E. Gur, R. C. Gur, and R. Verma. “Sex differences in the structural connectome of the human brain”. In: *Proceedings of the National Academy of Sciences* 111.2 (2014), pp. 823–828 (cit. on p. 50).

[32] I. Inza, P. Larrañaga, R. Blanco, and A. J. Cerrolaza. “Filter versus wrapper gene selection approaches in DNA microarray domains”. In: *Artificial intelligence in medicine* 31.2 (2004), pp. 91–103 (cit. on p. 19).

[33] B. Jeurissen, J.-D. Tournier, T. Dhollander, A. Connelly, and J. Sijbers. “Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data”. In: *NeuroImage* 103 (2014), pp. 411–426. ISSN: 1053-8119. URL: `http://www.sciencedirect.com/science/article/pii/S1053811914006442` (cit. on p. 14).

[34] B. Jeurissen, J.-D. Tournier, T. Dhollander, A. Connelly, and J. Sijbers. “Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data”. In: *NeuroImage* 103 (2014), pp. 411–426 (cit. on p. 35).

[35] B. Jie, M. Liu, D. Zhang, and D. Shen. “Sub-network kernels for measuring similarity of brain connectivity networks in disease diagnosis”. In: *IEEE Transactions on Image Processing* 27.5 (2018), pp. 2340–2353 (cit. on p. 66).

[36] D. Jones. *Diffusion MRI*. Oxford University Press, 2010. ISBN: 9780199708703. URL: `https://books.google.de/books?id=dbZCMePD52AC` (cit. on p. 12).

[37] D. K. Jones, T. R. Knösche, and R. Turner. “White matter integrity, fiber count, and other fallacies: The do’s and don’ts of diffusion MRI”. In: *NeuroImage* 73 (2013), pp. 239–254. ISSN: 1053-8119. URL: `http://www.sciencedirect.com/science/article/pii/S1053811912007306` (cit. on p. 64).

[38] R. A. Kanaan, M. Allin, M. Picchioni, G. J. Barker, E. Daly, S. S. Shergill, J. Woolley, and P. K. McGuire. “Gender differences in white matter microstructure”. In: *PloS one* 7.6 (2012), e38272 (cit. on p. 50).

[39] N. Kriege and P. Mutzel. “Subgraph Matching Kernels for Attributed Graphs”. In: *Proceedings of the 29th International Conference on Machine Learning, ICML 2012* 2 (June 2012) (cit. on p. 66).

[40] A. A. Loboda, M. N. Artyomov, and A. A. Sergushichev. “Solving Generalized Maximum-Weight Connected Subgraph Problem for Network Enrichment Analysis”. In: *Algorithms in Bioinformatics.* Ed. by M. Frith and C. N. Storm Pedersen. Cham: Springer International Publishing, 2016, pp. 210–221. ISBN: 978-3-319-43681-4 (cit. on pp. 20, 21, 29, 38, 40–42, 64).

[41] S. Mastrogiacomo, W. Dou, J. A. Jansen, and X. F. Walboomers. “Magnetic Resonance Imaging of Hard Tissues and Hard Tissue Engineered Bio-substitutes”. In: *Molecular Imaging and Biology* 21.6 (Dec. 2019), pp. 1003–1019. ISSN: 1860-2002. URL: `https://doi.org/10.1007/s11307-019-01345-2` (cit. on p. 5).

[42] W. McKinney et al. “pandas: a foundational Python library for data analysis and statistics”. In: *Python for High Performance and Scientific Computing* 14.9 (2011) (cit. on p. 29).

[43] D. W. McRobbie, E. A. Moore, M. J. Graves, and M. R. Prince. *MRI from Picture to Proton.* 2nd ed. Cambridge University Press, 2006 (cit. on p. 3).

[44] J. Muller, M. Alizadeh, F. B. Mohamed, J. Riley, J. J. Pearce, B. Trieu, T.-W. Liang, V. Romo, A. Sharan, and C. Wu. “Clinically applicable delineation of the pallidal sensorimotor region in patients with advanced Parkinson’s disease: study of probabilistic and deterministic tractography”. In: *Journal of neurosurgery* 131.5 (2018), pp. 1520–1531 (cit. on p. 15).

[45] B. Mwangi, T. S. Tian, and J. C. Soares. “A review of feature reduction techniques in neuroimaging”. In: *Neuroinformatics* 12.2 (2014), pp. 229–244 (cit. on p. 18).

[46] G. J. Parker, H. A. Haroon, and C. A. Wheeler-Kingshott. “A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements”. In: *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 18.2 (2003), pp. 242–254 (cit. on p. 34).

[47] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830 (cit. on p. 29).

[48] J. Ridgway. “Cardiovascular magnetic resonance physics for clinicians: Part I”. In: *Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance* 12 (Nov. 2010), p. 71 (cit. on p. 8).

[49] A. Romano, G. D'andrea, G. Minniti, L. Mastronardi, L. Ferrante, L. Fantozzi, and A. Bozzao. "Pre-surgical planning and MR-tractography utility in brain tumour resection". In: *European radiology* 19.12 (2009), p. 2798 (cit. on p. 14).

[50] M. Rubinov and O. Sporns. "Complex network measures of brain connectivity: uses and interpretations". In: *Neuroimage* 52.3 (2010), pp. 1059–1069 (cit. on p. 16).

[51] R. Sala-Llonch, D. Bartrés-Faz, and C. Junqué. "Reorganization of brain networks in aging: a review of functional connectivity studies". In: *Frontiers in psychology* 6 (2015), p. 663 (cit. on p. 1).

[52] A. Sarica, A. Cerasa, and A. Quattrone. "Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review". In: *Frontiers in Aging Neuroscience* 9 (2017), p. 329. ISSN: 1663-4365. URL: `https://www.frontiersin.org/article/10.3389/fnagi.2017.00329` (cit. on p. 2).

[53] S. Shi and F. Nathoo. "Feature Learning and Classification in Neuroimaging: Predicting Cognitive Impairment from Magnetic Resonance Imaging". In: *2018 4th International Conference on Big Data and Information Analytics (BigDIA)*. IEEE. 2018, pp. 1–5 (cit. on p. 18).

[54] A. Sifaoui, A. Abdelkrim, and M. Benrejeb. "On the Use of Neural Network as a Universal Approximator". In: *International Journal of Sciences* 2 (Aug. 2008), pp. 386–399 (cit. on p. 27).

[55] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly. "SIFT: Spherical-deconvolution informed filtering of tractograms". In: *Neuroimage* 67 (2013), pp. 298–312 (cit. on p. 35).

[56] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly. "The effects of SIFT on the reproducibility and biological accuracy of the structural connectome". In: *NeuroImage* 104 (2015), pp. 253–265. ISSN: 1053-8119. URL: `http://www.sciencedirect.com/science/article/pii/S1053811914008155` (cit. on p. 50).

[57] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly. "The effects of SIFT on the reproducibility and biological accuracy of the structural connectome". In: *Neuroimage* 104 (2015), pp. 253–265 (cit. on p. 64).

[58] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly. "Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information". In: *NeuroImage* 62.3 (2012), pp. 1924–1938. ISSN: 1053-8119. URL: `http://www.sciencedirect.com/science/article/pii/S1053811912005824` (cit. on p. 34).

[59] O. Sporns. "Brain connectivity". In: *Scholarpedia* 2.10 (2007). revision #91084, p. 4695 (cit. on p. 17).

[60] O. Sporns. “Graph theory methods: applications in brain networks”. In: *Dialogues in clinical neuroscience* 20.2 (2018), p. 111 (cit. on p. 16).

[61] O. Sporns and D. S. Bassett. “Editorial: New Trends in Connectomics”. In: *Network Neuroscience* 2.2 (2018), pp. 125–127. eprint: `https://doi.org/10.1162/netn_e_00052`. URL: `https://doi.org/10.1162/netn_e_00052` (cit. on p. 16).

[62] O. Sporns, G. Tononi, and R. Kötter. “The human connectome: a structural description of the human brain”. In: *PLoS Comput Biol* 1.4 (2005), e42 (cit. on p. 16).

[63] E. O. Stejskal and J. E. Tanner. “Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient”. In: *The journal of chemical physics* 42.1 (1965), pp. 288–292 (cit. on pp. 9, 11).

[64] J.-L. R. Stevens, P. Rudiger, and J. A. Bednar. “HoloViews: Building complex visualizations easily for reproducible science”. In: *Proceedings of the 14th Python in Science Conference*. 2015, pp. 61–69 (cit. on p. 48).

[65] B. Szalkai, B. Varga, and V. Grolmusz. “Graph theoretical analysis reveals: Women’s brains are better connected than men’s”. In: *PLoS One* 10.7 (2015), e0130045 (cit. on p. 50).

[66] P. Tahmasebi, S. Kamrava, T. Bai, and M. Sahimi. “Machine learning in geo- and environmental sciences: From small to large scale”. In: *Advances in Water Resources* 142 (2020), p. 103619. ISSN: 0309-1708. URL: `http://www.sciencedirect.com/science/article/pii/S0309170820301366` (cit. on p. 25).

[67] Y. Taki, B. Thyreau, S. Kinomura, K. Sato, R. Goto, R. Kawashima, and H. Fukuda. “Correlations among brain gray matter volumes, age, gender, and hemisphere in healthy individuals”. In: *PloS one* 6.7 (2011), e22734 (cit. on p. 50).

[68] D. G. Taylor and M. C. Bushell. “The spatial mapping of translational diffusion coefficients by the NMR imaging technique”. In: *Physics in Medicine and Biology* 30.4 (Apr. 1985), pp. 345–349. URL: `https://doi.org/10.1088%2F0031-9155%2F30%2F4%2F009` (cit. on p. 9).

[69] K. D. Toennies. *Guide to medical image analysis*. Springer (cit. on p. 9).

[70] J. Tohka, E. Moradi, H. Huttunen, A. D. N. Initiative, et al. “Comparison of feature selection techniques in machine learning for anatomical brain MRI in dementia”. In: *Neuroinformatics* 14.3 (2016), pp. 279–296 (cit. on p. 18).

[71] J. D. Tournier, F. Calamante, and A. Connelly. “Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions”. In: *Proceedings of the international society for magnetic resonance in medicine*. Vol. 1670. Ismrm. 2010 (cit. on p. 35).

[72] J.-D. Tournier, F. Calamante, D. G. Gadian, and A. Connelly. "Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution". In: *Neuroimage* 23.3 (2004), pp. 1176–1185 (cit. on pp. 13, 14).

[73] J.-D. Tournier, R. Smith, D. Raffelt, R. Tabbara, T. Dhollander, M. Pietsch, D. Christiaens, B. Jeurissen, C.-H. Yeh, and A. Connelly. "MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation". In: *NeuroImage* 202 (2019), p. 116137 (cit. on pp. 29, 45).

[74] J.-D. Tournier, F. Calamante, D. G. Gadian, and A. Connelly. "Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution". In: *NeuroImage* 23.3 (2004), pp. 1176–1185. ISSN: 1053-8119. URL: `http://www.sciencedirect.com/science/article/pii/S1053811904004100` (cit. on p. 14).

[75] D. C. Van Essen and M. F. Glasser. "The human connectome project: progress and prospects". In: *Cerebrum: the Dana forum on brain science*. Vol. 2016. Dana Foundation. 2016 (cit. on p. 1).

[76] D. C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S. W. Curtiss, et al. "The Human Connectome Project: a data acquisition perspective". In: *Neuroimage* 62.4 (2012), pp. 2222–2231 (cit. on p. 31).

[77] S. Vieira, W. H. Pinaya, and A. Mechelli. "Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications". In: *Neuroscience & Biobehavioral Reviews* 74 (2017), pp. 58–75 (cit. on p. 26).

[78] J. G. White, E. Southgate, J. N. Thomson, and S. Brenner. "The structure of the nervous system of the nematode Caenorhabditis elegans". In: *Philos Trans R Soc Lond B Biol Sci* 314.1165 (1986), pp. 1–340 (cit. on p. 17).

[79] R. K. W. Wong, T. C. M. Lee, D. Paul, J. Peng, and A. Disease Neuroimaging Initiative. "Fiber direction estimation, smoothing and tracking in diffusion MRI". In: *Ann. Appl. Stat.* 10.3 (Sept. 2016), pp. 1137–1156. URL: `https://doi.org/10.1214/15-AOAS880` (cit. on p. 9).

[80] T. Yamamoto, H. Bannai, M. Nagasaki, and S. Miyano. "Better decomposition heuristics for the maximum-weight connected graph problem using betweenness centrality". In: *International Conference on Discovery Science*. Springer. 2009, pp. 465–472 (cit. on p. 20).

[81] C.-H. Yeh, D. K. Jones, X. Liang, M. Descoteaux, and A. Connelly. "Mapping Structural Connectivity Using Diffusion MRI: Challenges and Opportunities". In: *Journal of Magnetic Resonance Imaging* (2020) (cit. on pp. 48, 50, 64).

Appendix

In this section the lookup table (fs_default.txt) for relevant gray matter parcellations in *MRtrix3* is given. In this full name column of this section the short form ‘ctx’ stands for cortical region, ‘lh’ for left hemipshere and ‘rh’ for right hemisphere. The ROIs which do not have the short form ‘ctx’ in their name correspond to subcortical gray matter regions.

Label Number	**Label Name**
1	ctx-lh-bankssts
2	ctx-lh-caudalanteriorcingulate
3	ctx-lh-caudalmiddlefrontal
4	ctx-lh-cuneus
5	ctx-lh-entorhinal
6	ctx-lh-fusiform
7	ctx-lh-inferiorparietal
8	ctx-lh-inferiortemporal
9	ctx-lh-isthmuscingulate
10	ctx-lh-lateraloccipital
11	ctx-lh-lateralorbitofrontal
12	ctx-lh-lingual
13	ctx-lh-medialorbitofrontal
14	ctx-lh-middletemporal
15	ctx-lh-parahippocampal
16	ctx-lh-paracentral
17	ctx-lh-parsopercularis
18	ctx-lh-parsorbitalis
19	ctx-lh-parstriangularis
20	ctx-lh-pericalcarine
21	ctx-lh-postcentral
22	ctx-lh-posteriorcingulate
23	ctx-lh-precentral
24	ctx-lh-precuneus

Label Number	**Label Name**
25	ctx-lh-rostralanteriorcingulate
26	ctx-lh-rostralmiddlefrontal
27	ctx-lh-superiorfrontal
28	ctx-lh-superiorparietal
29	ctx-lh-superiortemporal
30	ctx-lh-supramarginal
31	ctx-lh-frontalpole
32	ctx-lh-temporalpole
33	ctx-lh-transversetemporal
34	ctx-lh-insula
35	Left-Cerebellum-Cortex
36	Left-Thalamus
36	Left-Thalamus-Proper
37	Left-Caudate
38	Left-Putamen
39	Left-Pallidum
40	Left-Hippocampus
41	Left-Amygdala
42	Left-Accumbens-area
43	Right-Thalamus
43	Right-Thalamus-Proper
44	Right-Caudate
45	Right-Putamen
46	Right-Pallidum
47	Right-Hippocampus
48	Right-Amygdala
49	Right-Accumbens-area
50	ctx-rh-bankssts
51	ctx-rh-caudalanteriorcingulate
52	ctx-rh-caudalmiddlefrontal
53	ctx-rh-cuneus
54	ctx-rh-entorhinal
55	ctx-rh-fusiform
56	ctx-rh-inferiorparietal
57	ctx-rh-inferiortemporal
58	ctx-rh-isthmuscingulate
59	ctx-rh-lateraloccipital

60	ctx-rh-lateralorbitofrontal
Label Number	**Label Name**
61	ctx-rh-lingual
62	ctx-rh-medialorbitofrontal
63	ctx-rh-middletemporal
64	ctx-rh-parahippocampal
65	ctx-rh-paracentral
66	ctx-rh-parsopercularis
67	ctx-rh-parsorbitalis
68	ctx-rh-parstriangularis
69	ctx-rh-pericalcarine
70	ctx-rh-postcentral
71	ctx-rh-posteriorcingulate
72	ctx-rh-precentral
73	ctx-rh-precuneus
74	ctx-rh-rostralanteriorcingulate
75	ctx-rh-rostralmiddlefrontal
76	ctx-rh-superiorfrontal
77	ctx-rh-superiorparietal
78	ctx-rh-superiortemporal
79	ctx-rh-supramarginal
80	ctx-rh-frontalpole
81	ctx-rh-temporalpole
82	ctx-rh-transversetemporal
83	ctx-rh-insula
84	Right-Cerebellum-Cortex

Acknowledgements

Completing my master's thesis during an ongoing pandemic was a big challenge, both personally and professionally. It was through the constant support of my supervisors, family and friends that I have been able to complete this thesis. First of all, I am very grateful to Prof. Dr. Thomas Schultz for the support and guidance during the course of this thesis. His insights have been valuable for developing a consolidated research plan and viewing the research problem critically. He also made it possible for me to receive a fellowship from the Deutscher Akademiker Austauchdienst (DAAD). I am also very thankful to him for delivering interesting courses that sparked my research interests and stimulated me to research on this topic.

Secondly, I would like to thank Mohammad Khatami for being my day-to-day supervisor and being patient with all my questions. He gave me constant feedback and was immensely helpful in solving implementational problems. I am also indebted to Dr. Regina Wehler whose Master Thesis Research made it easy for me to get started with my research topic.

I am also thankful to the DAAD and the German Foreign Office (AA) for financially supporting my master's thesis. The financial support made it possible for me to maintain my focus during my project.

Data for the project were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil)funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Bonn, November 16, 2020

Location, Date

Signature

Declaration of Authorship

I herewith certify that this material is my own work, that I used only those sources and resources referred to in the thesis, and that I have identified citations as such. It has not been submitted, either in part or whole, for a degree at this or any other university.

Bonn, November 16, 2020

Location, Date

Signature