University of Bonn

Master's thesis for obtaining the academic degree
„Master of Science (M.Sc.)“

Extracting most predictive subgraphs from models of human brain connectivity

Author:
Shreya Kapoor

First Examiner:
Prof. Dr. Thomas Schultz

Second Examiner:
Prof. Dr. Holger Fröhlich

Advisor:
Mohammad Khatami

Submitted: October 17, 2020

Declaration of Authorship

I declare that the work presented here is original and the result of my own investigations. Formulations and ideas taken from other sources are cited as such. It has not been submitted, either in part or whole, for a degree at this or any other university.

Location, Date

Signature

Acknowledgements

funding bodies supervisors professor laboratory assistants librarians colleagues editors/proofreaders (a requirement at some universities) classmate research participants (e.g. people who completed a survey to help you gather data)
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 greatful to Prof. Dr. Thomas Schultz for the support and guidance during the course this thesis. His insights have been very 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 DAAD (Deutscher Akademiker Austauchdienst). Without his nomination, I would not have received a fellowship position. I am also very thankful to him for delivering interesting and elucidate courses that sparked my research interests and propelled me to research on this topic.

Secondly, I would like to thank Mohammad Khatami for being my day-to-day supervisor and being very patient with all my questions. He gave me constant feedback and was immensely helpful solve 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.

Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657)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.

Location, Date

Signature

Abstract

Advances in non-invasive Neuroimaging modalities has lead to a growing interest in investigating neural underpinnings of human cognition and neurological disorders. Analysis of human brain connectivity or the human connectome has hence become an active area of research. On one hand, Computational methods, especially the ones based on graph theory have played a significant role in understanding the topological organisation of brain networks. On the other hand, machine learning classifiers 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 in order to obtain discriminative subnetworks in the brain, which make the classification results more interpretable. The novel component of this thesis is to reduce the graphical network derived from structural brain connectivity by approximating the solution to the Maximum Edge Weight Subgraph (MEWS) problem that falls into the class of NP-hard problems. The comprehensive analysis pipeline includes the comparison of feature selection based techniques based on statistical analysis, and the MEWS based method. Experimental results from this technique show (to be filled after discussion)

Contents

List of Figures

List of Tables

List of Algorithms

1 Introduction

One of the greatest scientific challenges in the 21^{st} century is to understand the structure (composition) and function of the human brain. How do intricate network of cells (neurons) organize themselves at different scales and give rise to behaviour, emotion and intelligence has remained an open question in neuroscience since ages. In recent years, there has been a surge of interest in understanding the human brain in health and disease, and on a more philosophical level, trying to understand 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, computational memory resources has lead to development of imaging standards and made wide range studies possible. Increasing efficiency and accuracy of non-invasive imaging modalities such as Magnetic Resonance Imaging (MRI), CT and PET have contributed significantly towards the human brain in health and disease. Among these, MRI is one of the most preferred modalities due to the diversity of information that can be generated using one scanner only. Each of the four types structural MRI, Task-activated functional MRI, Diffusion MRI and Resting state fMRI can be used to carry out different types of research (**van2016human**).

Excitement from investigators has brought about the progress of ambitious projects such as the Human Connectome Project launched in 2009 (**van2016human**). It explores the research area known as 'Connectomics' (previously known as hodology), the study of brain's structural and functional networks. It aims at creating a map of the brain divided into functionally distinct areas known as 'parcels', understand how these areas are connected to each other and how they contribute to behaviour. This framework is also important to understand what goes wrong during Neurological disorders **sala2015reorganization**.

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 functional Magnetic Resonance Imaging (fMRI). Each of these connectivities have their own importances but DTI based structural connectivity is of special importance in medical applications as it provides insight into anatomical connections in the brain of individual patients, can aid precision medicine

(**cociu2017multimodal**).

Machine Learning has become an indispensable tool in Neuroscience/for medical research due to the fact that the number of subjects in the study is often much lesser than the number of features in the data. It has often been used to make predictions based on MRI scans and classify diseased states of patients. However, for the analysis of brain connectivity there is a need for the ML algorithm to understand the topological properties of brain organization. It is also important to realize that the brain connectivity can be seen as a dense graph and there are not too many graph based machine learning classifiers. Further, even if these classifiers are able to work with graphs, they suffer from a lack of interpretability.

10.3389/fnagi.2017.00329

2212-

2 Background

Non-invasive neuroimaging techniques have become indispensable tools to understand cognitive processes and neurological disease pathology. Magnetic Resonance Imaging (MRI) is one of the most widely used neuroimaging modalities preferred by clinicians due to the safety of its acquisition procedure; it does not produce any ionizing radiation like X-ray or CT scans. It is of high utility when required to look at soft tissues such as those in the nervous system.

2.1 Magnetic Resonance Imaging

MRI is a non-invasive imaging technology that produces 3-dimensional detailed anatomical images(**mcrobbie_moore_graves_prince_2006**). This technique is based on the Nuclear Magnetic Resonance(NMR) Imaging principle. NMR a physical phenomenon in which nuclei respond to a combination of a constant and weak oscillating magnetic field by producing a signal at the resonant frequency of the nucleus. The difference between the magnetic properties of different tissue types subjected to an external magnetic field is used to generate images. Such electromagnetic properties are used to form images since our body is 65% water (H_2O), has a dipole moment and the hydrogen nuclei act like little magnets due to the fact that hydrogen nuclei have an odd number of protons and non-zero net spin.

In normal conditions (Fig. 2.1 part a), i.e. in the absence of an external magnetic field the spins of the hydrogen nuclei are randomly oriented and 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 and transverse magnetization as the one perpendicular to it. This can be represented using the equation

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

where M is the magnetization, μ is the magnetic permeability or the inertia of a material to get magnetized and B_0 is the external magnetic field. As per fig. 2.1 part a, it can be seen that the direction of B_0 defines the coordinate system

for the experiment. The transverse plane is defined as the plane perpendicular to the direction of B_0 whereas, the longitudinal magnetization is termed as the magnetization experienced in the direction of the static magnetic field. 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 exernal field.

When a radio wave is added to the static field using a radio frequency coil (RF coil) at the Larmour frequency (eqn. 2.2), termed as B_1 in 2.1 part a, the direction of the magnetization vector gets altered i.e. 'flipped' or 'tipped' out of alignment with B_0 towards B_1 with the angle of rotation being termed as the flip angle. The RF pulse causes exerts torque. Mathematically, the torque can be expressed as

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

where $\vec{m}$ represents 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 the net magnetization gets flipped (M_{xy}) onto the transverse plane. The magnetization M_{xy} precesses

The net magnetization does not precess until an external force disturbs its equilibrium position of alignment with the static magnetic field. Thus, when the magnetization precesses, magnetic resonance is said to occur. The magnetization has an effect in which the frequency is proportional to the applied static magnetic field.

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 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. **PhysRev.70.460** 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 part (b) with T1 relaxation being termed as the process by which the net magnetization aligns itself with the direction of the original magnetic field. The T2 constant measures the decay of the transverse component of the magnetization(Fig. 2.1 part (c)). The value of T2 reflects the time required for the magentization to

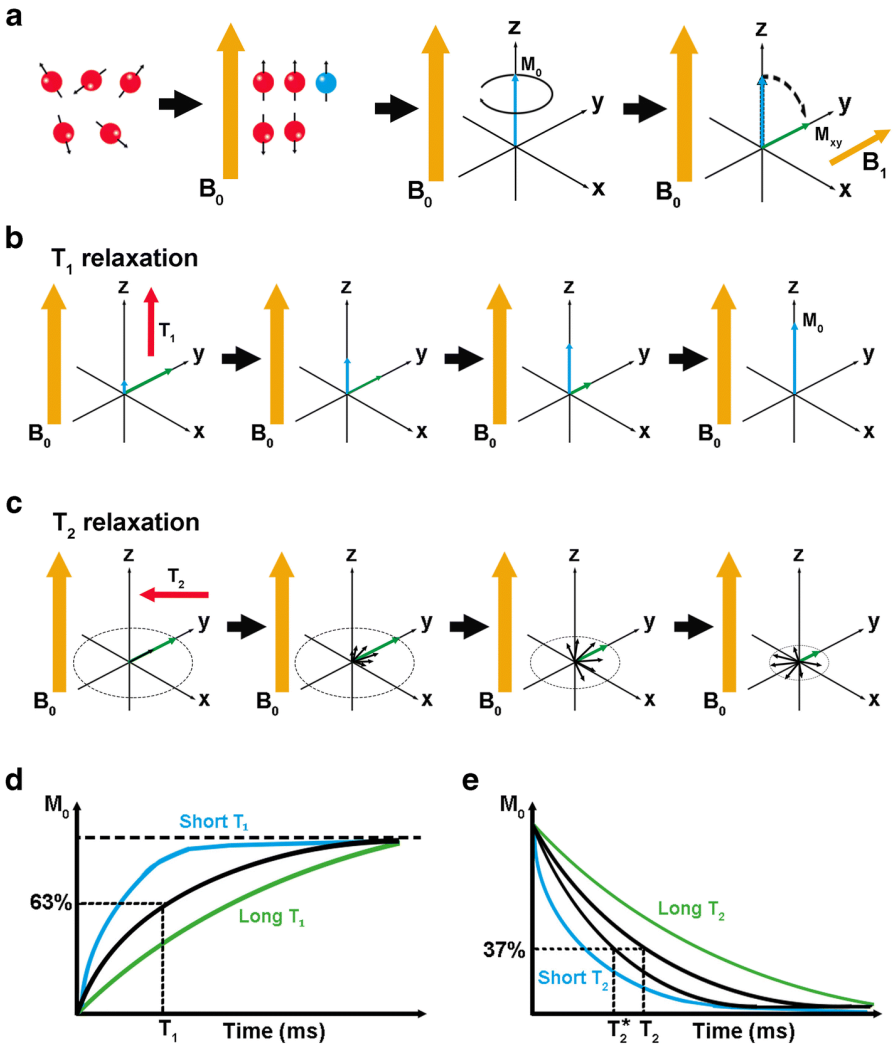

Figure 2.1: figure from **Mastrogiacomo2019**. (a) At first the magnetic moments of the individual nuclei are oriented randomly and cancel out each other, in the presence of an external field B_0 they gain a net magnetization, with the presence of the external RF pulse the phase of the net magnetization is changed and magnetic resonance is said to occur. (b) Describes process of T1 relaxation. The value of the T1 constant is the time required to achieve 63% of the longitudinal magnetization. (b) Describes T2 relaxation. The time constant T2 is the time required for the magnetization to fall to 1/e of its value.

fall to $\frac{1}{e}$ or 37% of its original value, reflected in fig. 2.1 part e. Since inside the MRI scanner there can be imhomogneties, another constant $T2^*$ is measured as the T2 time observed while taking the recording. During the resonance phenomena, the magnetization has components in different directions which can be expressed in terms of the time constants and current time

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

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

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

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 the equation 2.6 it is clear that the T1 constant measure the time required to gain the original magnetization, i.e. when $t = T1$ $M_z = M_0$ and when $t = T2$, $M_{xy} = \frac{M_0}{\cos \omega T2}$

T1 weighting is generated on the basis of controlling the repetition time (TR) in a sequence. The TR is defined as the time taken between sucessive excitations of the same region. Using a small TR causes the net magntizations aligned along the fileds to not fully recover. Different brain tissues have different T1 values. Short T1 times are seen in fat molecules because the complex structure of the saturated molecules leads to flexions and roations which might occur at the Larmour frequency, thus making it easier to return to the initial state, while longer T1 times is observed in comparatively freely diffusing mediums sch as cerebrospinal fluid (CSF). The fat molecules appear bright on the T1-weighted images while liquids such as cerebrospinal fluid appear dark due to longer T1 times.

Similarly, T2 weighting of an image is generated on the basis of echo time (TE). The echo time measures time between excitation and signal measurement. A well adjusted, long TE gives rise to contrast of signals from different images. Fat molecules appear lighter on T2-weighted images due to short T2 time while the cerebro-spinal fluid appears bright since the CSF has long T2 times.

2.1.1 Image formation

An MRI scan needs to represent the 3D nature of the body part under study. Inside an MRI scanner, the gradient coils are hence used to alter the magnetic field along the different spatial directions so that different slices of the subject's body resonate at different frequencies. The signal is then received by receiver coils that detect only the transverse magnetization. The signal needs spatial encoding

since information about the millions of voxels in the brain has to be maintained.

For the image formation, we know that the magnetic gradients are applied in the x,y,z directions (in order to get the 3D image). This formalism results in each voxel possessing a different Larmour frequency, termed as spatial encoding.

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

According to this equation, it is prominent 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 2-dimensional image reconstruction has been explained in the figure 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 pahse 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) and 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 but it is not necessary that the locations in both the images are exactly the same. This method is then extrapolated in 3D in order to obtain the image of the whole organ such as the brain. Any weighting, T1 or T2 can then be given in order to generate the image.

Another important concept is that of Field of View (FOV) 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), $\Delta k = 1/FOV$.

2.1.2 Diffusion MRI

Diffusion Magnetic resonance imaging 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 (**ghosh2015survey**). One of the most important applications of dMRI is to map the white-matter fiber tracts in the brain. As compared to other MRI

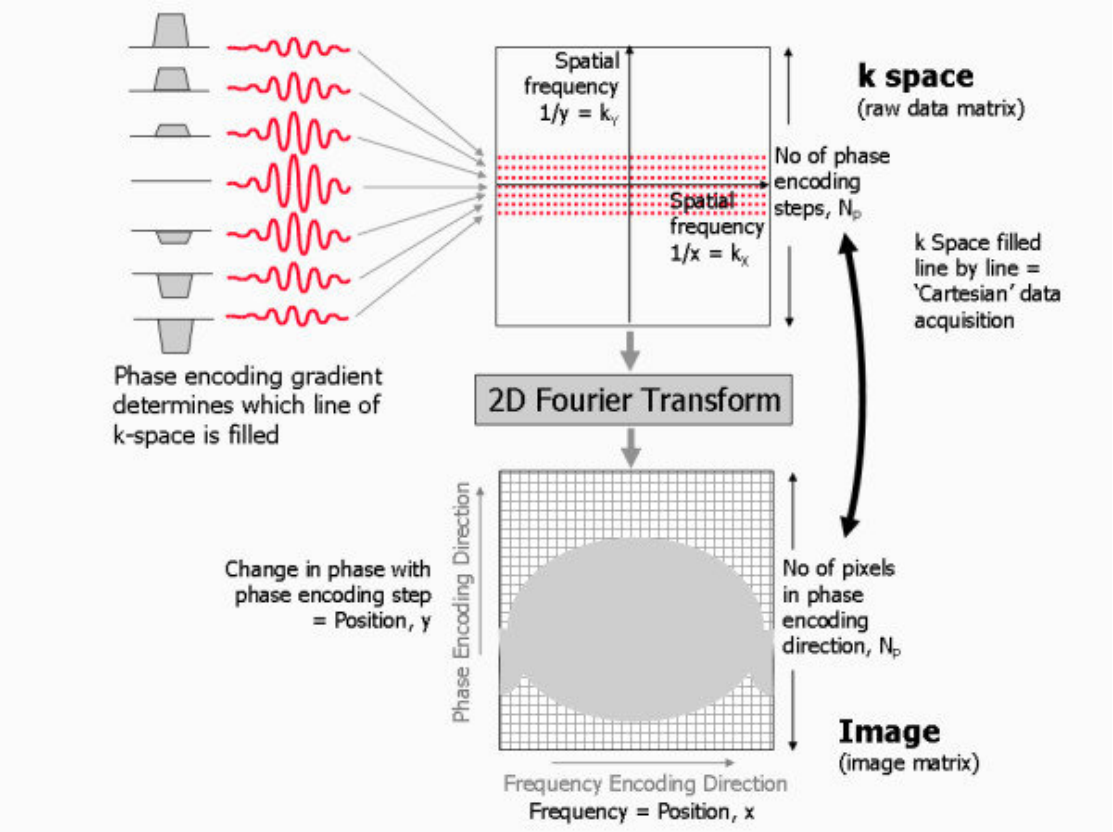

Figure 2.2: Image from **MRIrecon** summarising the reconstruction of an encoded MRI signal in 2D, after a slice has been selected using the gradient G_z. The phase is encoded along the y direction. Each phase encoding step is used to populate the k-space and the number of such steps determines the number of pixels in the y-direction. The frequency is encoded along the x direction. The relationship between the k-space points and the points in reconstructed image space is such that $k_x = 1/x$ and $k_y = 1/y$. Each line parallel to the k_x axis corresponds to a separate MRI signal. Each line parallel to the k_y axis corresponds to the amplitude and duration of the phase encoding direction at each phase encoding step.

techniques such as functional Magnetic Resonance Imaging it does not suffer from an issue of low resolution and low signal-to-noise ratio (**wong2016**).

"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. they can move in any direction with equal probability or exhibit random walk behavior explained by Brownian motion (**Brogioli_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. Water molecules in the extracellular environment hence experience relatively free diffusion and the ones in the intra-cellular environment experience restricted diffusion (**toennies2017guide**). This diffusion anisotropy is encoded into the MRI signal using spatial and temporal variation(gradients) in the magnetic field (the alignment of the molecules which have different diffusion hence becomes different). The MRI signal is said to be "Diffusion Weighted" due to the signal attenuation introduced by the magnetic field gradients.

2.1.2.1 Diffusion Weighted Imaging

In this type of Diffusion imaging, 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 presumption that diffusion varies with pathology i.e. differences in diffusion can also highlight differences in structure and function (**Taylor_1985**).

One of the most popular ways to give images Diffusion weighting is through using single-shot-spin-ehco (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 (**stejskal1965spin**) in the mid 1960's, which resulted in much improved sensitivity to diffusion in comparison to the steady state gradients used previously. They solved the Bloch-Torrey partial differential equations for a symmetric pair of pulsed gradients and obtained the well-known Stejskal-Tanner formula

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

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 the attenuation of the diffusion MRI signal by the diffusion gradients is represented in figure 2.3. The application of the Diffusion gradient results in a spatial encoding as the Larmour frequency is dependent on the net magnetic

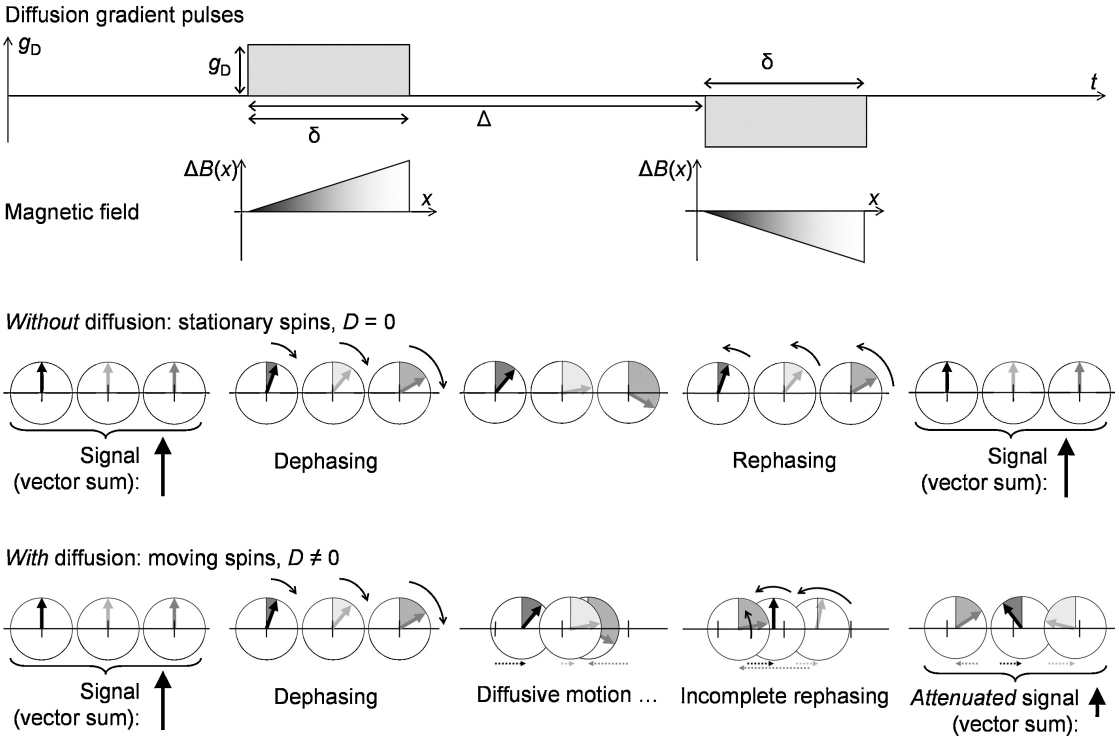

Figure 2.3: Image from **technical_aspects_dmri** elucidating the technical aspects of Diffusion Weighted MRI. The RF pulse causes a phase change of the magnetization. Introducing the diffusion gradient adds spatial dependence of phase shift of the individual magnetic moments. The molecules that have restricted diffusion remain along do not experience the effect of the diffusion gradient, any phase change from the first gradient is reversed by the second, and tend to relax to their equilibrium state. However, 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 phase shift is then manifested in terms of signal attenuation. The degree of signal attenuation depends on multiple factors as shown in the following equation: SI0 is the signal intensity of the T2-weighted image with no diffusion gradient applied, b is the degree of diffusion weighting (b value), and D is the diffusion coefficient

field (equation 2.7). In the figure 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 of $\Delta w(x)$, 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 Diffusion gradients) while with Diffusion weighting there is a signal attenuation explained by equation 2.8.

The where the b value of the sequence 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. Also, the b values need to be adjusted in order to obtain optimal signal to noise ratio (SNR)

2.1.2.2 Diffusion Tensor Model

Diffusion Tensor Imaging (DTI) is a new type of imaging technique that relies on a tensor model to measure the diffusion among voxels. Instead of attributing diffusion inside a voxel by using 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 and hence able to trace complex fiber tracts in the brain. (**jones2010diffusion**)

Diffusion Tensor Imaging is a novel technology, an *in-vivo* application of DWI that 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 defects in white matter. 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 or the 'connectome' using DTI does not only help to

understand the pathophysiological effect of brain disorders but also the structure of functional networks.

In this type of imaging, each voxel is associated with a 3x3 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 corresponding eigenvectors which represent the directions of Diffusion along the voxel, the voxels are usually 1 mm3 in size and often constitute components of more than one cell within them.

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

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

$$\frac{S}{S_0} = \exp^{-bg^T Dg} \tag{2.10}$$

where g^T is a 3x1 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". Diffusion in complex environments cannot be explained by using diffusion gradients in one direction only. In fig 2.3 of the Diffusion Weighted MRI, 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 hindered/interacting with cell membranes/extracellular environments.

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 (R,G, B 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.

There are three important quantities that can easily be derived from the diffusion tensor. These help to analyze the nature of differences in diffusion along the various voxels of the image. The three quantities are mean diffusivity, diffusion anisotropy and fractional anisotropy. The trace of the Diffusion tensor is used as Mean diffusivity. While the diffusion anisotropy exhibits 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 (an indication of more white matter?

The fractional anisotropy determines a sort of 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,D2 the diffusion coefficients along xx,yy,zz)

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

$$FA = \sqrt{\frac{3}{2}} \frac{\sqrt{(D1 - D_{mean})^2 + (D2 - D_{mean})^2 + (D3 - D_{mean})^2}}{\sqrt{D1^2 + D2^2 + D3^2}} \tag{2.11}$$

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, its fails in regions containing populations of fiber orientations that have different orientations. It assumes that all white matter bundles in the brain have similar diffusion characteristics and attributes diffusion anisotropy to partial volume effects **tournier2004direct**. 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 in DTI, 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 probelms of kissing fibers/crossing fibers etc.

These rely on the concept of an fiber orientation distribution function (fODF) An fODF is a symmetric probability distribution function describing the distribution of fiber orientations

$$F(\Theta, \phi) = \sum_{k=1}^{K} w_k \delta_{\theta_k, \phi_k}(\Theta, \phi) \qquad \Theta \in [0, \pi], \phi \in [0, 2\pi] \tag{2.12}$$

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.

High angular resolution imaging (HARDI) techniques enable the detection of multi-modal diffusion signals. Q -ball imaging is one such technique but has its own limitations **TOURNIER20041176**.

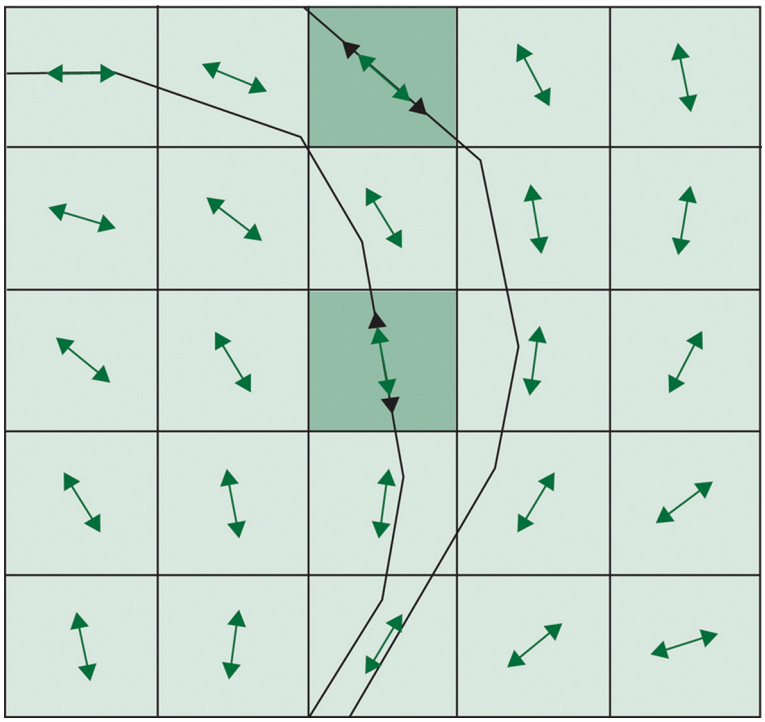

Figure 2.4: Deterministic tractography from

The DWI signal is modelled as the spherical convolution (i.e. convolution in spherical coordinates) of the response function and the fODF. Spherical deconvolution is the other way round where fODF can be extracted.

Constrained Spherical Deconvolution (CSD) has gained recent attention as a method to extract white matter (WM) fibre orientation distributions. It is used to obtain fODF estimates and fibre tractograms **JEURISSEN2014411**.

2.1.3 Tractography

Fiber tractography is a technique used to graphically construct 3 dimensional representations 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 **romano2009pre**.

Tractography methods often use DTI in order to delineate the streamlines along the white matter tracts by getting the information about their progression from one voxel to the next. This methodology 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.

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 then the algorithm proceeds tracking along the maximum diffusion direction of the second voxel and terminates when the condition is not met. Several other conditions can also be imposed on the basis tracking technique on the basis of investigator specific techniques. The deterministic tractography does not take into account any random or systematic errors that arise during signal acquisition, recording and transmission.

2.1.4 Analyzing the brain as a graph

The human brain is considered to be one of the most intricate biological systems. It has been well established that this complex system can be modelled as a network composed of various dynamically interacting elements. It can hence be well represented as a graph for computational analysis. The networks perspective has gained tractable attention in Neuroscience and given rise to an ever expanding field of Network Neuroscience. Brain networks can vary at scale (from molecular interactions to cognition. With different interactions at each level, graph theory finds its applications in this field **sporns2018graph**.

Graph theory can be defined as a branch of discrete mathematics that deals with the geometric analysis of graphs. Graph theory is useful in the analysis of biological networks since biological systems are dynamic and the interaction between the elements of the network gives rise to emergent properties. These properties can only be deciphered by comprehending the behavior with a systems approach. Graph theoretic methods can give comprehensive information about brain network topology, archictecture, structure and function. This has given rise to the field of connectomics **sporns2005human** for analysis of the structural and functional brain connectivity. The strucutral networks usually represent the anatomy of the brain and are temporally stable (ignoring 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 and causal relationship various factors

2.1.5 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 (**bassett2017network**). The connectome of 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 with evidence about the hodological view. The research in the field has expanded rapidly due to the increasing interest in understanding the brain as a dynamic system, and the belief that the connections in the brain are what give it its capabilities and functionality. **network_neuroscience_editorial**

The first complete structural connectome to be mapped at the synaptic level was published in 1986, belongs to the species C. elegans. It was reconstructed using electron micrographs of serial section and is a network that possess three hundred neurons and roughly seven thousand connections. It still took decades to claim the biological plausibility of all the connections in the reconstructed connectome**elegans**, this leads to the fact that the connectome of the human brain cannot be mapped using the manual labour intensive methodology used by **white1986structure**. Mapping the human brain manually hence seems implausible. The human brain contains roughly the same number of neurons as the stars in the Milky way galaxy and about 10^{15} inter neuronal connections (synapses)**fornito2015connectomics**. This calls for the need to computationally determine brain connectivity from brain scans. Diffusion MRI can be used for sturcutral connectome construction while functional MRI is the preferred modality for functional connectome construction. The Human Connectome Project 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.

2.1.6 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 derived from whole brain tractography.In organisms such as humans with complex nervous systems, 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**).

Usually in Neuroimaging studies the connectivity is analyzed at the macroscopic scale, i.e. the running of the fibers between different brain regions. From section **??** it can be inferred that the streamlines traced using the tractography algorithms represent the structural connections between different Regions of Interest(ROIs) of the brain seen as brain nodes. The connections are usually represented using bio-physical parameters such as mean FA (mean fractional anisotropy), number of streamlines between the two nodes and the length of the streamlines connecting the two nodes.

2.1.7 Functional Brain Connectivity

Funtional connectivity can be defined as the numerical analysis of brain regions engaged in a task. Any two brain regions are considered to be functionally connected if there exists a temporal correlation in activities recorded from them. The assumption in the use of functional connectivity is that two temporally coupled events in the brain could imply correlation in their association i.e. they are more associated to each other than to all those which are not temporally correlated.

The brain's functional connectivity can be analysed using fMRI data. It can be seen as a sort of dynamic connectivity.

2.2 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. Hoewever, most dimensionality reduction methods such as Principal Component Analysis (PCA) rely on transformations (often non-linear) of the original data and this leads to a lack of interpretability. Due to this reason it is important to consider feature selection methods where inference about what the classifier actually learns can be made (**shi2018feature**).

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 *on the fly* in which 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. These have been increasingly used in Neuroimaging studies **tohka2016comparison**.

2.2.1 Filter methods

Filters form one of the simplest methods for feature selection. They usually serve as a pre-processing step for the classification and are model independent. Hoeever, there is a caveat when using these type of features for classification studies that each feature is considered independently and their cumulative effect is not taken into consideration for the cla Some of the common techniques used to determine the feature importance is the use of coefficients such as the fscore and t-test, these are also used in Neuroimaging applications.

The fscore can be calculated as

$$F = \frac{(\overline{x_1} - \overline{x})^2 + (\overline{x_2} - \overline{x})^2}{\frac{\sum_{i=1}^{n1}(x_{1,i} - \overline{x_1})^2}{n_1 - 1} + \frac{\sum_{i=1}^{n1}(x_{2,i} - \overline{x_2})^2}{n_2 - 1}} \tag{2.13}$$

where $\overline{x}$ represents the average for all feature values, $\overline{x_1}$ the feature average for for the first class, n_1 represents the number of samples for the first class and other variables follow respectively.

A simple t-test can also be computed and then the features can be filtered based on the t-statistic or the p-value to test the significance of the differences between the feature values for two different groups. Based on **inza2004filter**, the t statistic is calculated as:

$$t = \frac{|\overline{x_1} - \overline{x_2}|}{\sqrt{0.5(\sigma_1^2 + \sigma_2^2)}} \tag{2.14}$$

The features can be ranked and the features with the highest t-test scores can be used to perform the classification. Numerical thresholds can also be imposed if suited to the nature of the problem.

2.2.2 Maximum Weight Subgraph

In the above section it was presented that the brain can be analyzed as a graph.

The definition of the maximum weight connected subgraph can be given as

$$\Omega(\tilde{G}) = \sum_{vin\tilde{V}} w_v + \sum_{e \in \tilde{E}} w_e \longrightarrow max \tag{2.15}$$

Such a problem is formulated using a Mixed INteger programming formulation.

The representation is as follows for the subgraph. the variables are

- Binary variable $y_v = 1$ iff $v \in V$ and $v \in \tilde{V}$
- Binary variable $w_e = 1$ iff $e \in E$ and $e \in \tilde{E}$

These variables can represent a valid subgraph iff for the constraints be satisfied.

$$w_e \leq y_v \forall v \in, e \in \delta_v$$

For the non-linear formulation the graph can be considered as set of binary variables and continuous variables

- 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 traversal from the root of the arborescence to the vertex v goes through n vertices. The d_v value can be arbitrary if the vertex v does not belong to the optimal solution.

To ensure the validity of the arborescence the following constraints were introduced.

$$\sum_{v \in V} r_v = 1 \tag{2.16}$$

$$1 \leq d_v \leq n \qquad \forall v \in V \tag{2.17}$$

$$\sum_{(u,v) \in A} x_{uv} + r_v = y_v \qquad \forall v \in V \tag{2.18}$$

$$x_{uv} + x_{vu} \leq w_e \qquad \forall e = (v,u) \in A \tag{2.19}$$

$$d_v r_v = r_v \tag{2.20}$$

Linearization of the above inequalities

$$d_v + n r_v \leq n \qquad \forall v \in V \tag{2.21}$$

$$n + d_u - d_v \geq (n+1) x_{vu} \qquad \forall (v,u) \in A \tag{2.22}$$

$$n + d_v - d_u \geq (n+1) x_{vu} \qquad \forall (v,u) \in A \tag{2.23}$$

$$\tag{2.24}$$

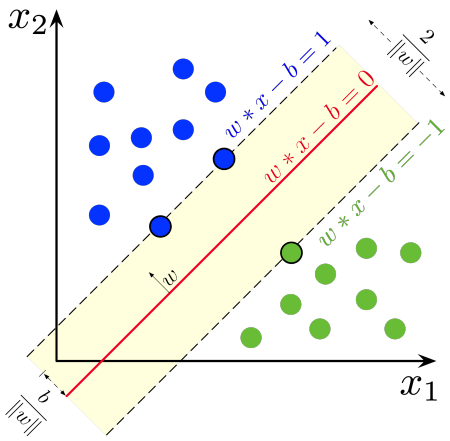

Figure 2.5: Caption

2.3 Classification

2.3.1 Support Vector Classifiers

Support Vector Classifiers (SVC) are classical machine learning algorithms based on Support Vector machines (SVMs). They are important since they have attained signifcant accuracy with different types of tasks such as a handwritten digits recognition, face detection in images, and text categorization**burges1998a**. 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 figure **??** the red line represents the optimally separating hyperplane that satisfies the equation $\sum w_i x_i - b = 0$ 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.

$$y_i(\sum_{j=1}^{d} w_j x_{ij} + b) - 1 \geq 0 \qquad \forall i \in 1...n \tag{2.25}$$

Here, the weights represent the parameters the model learns in order to satisfy the

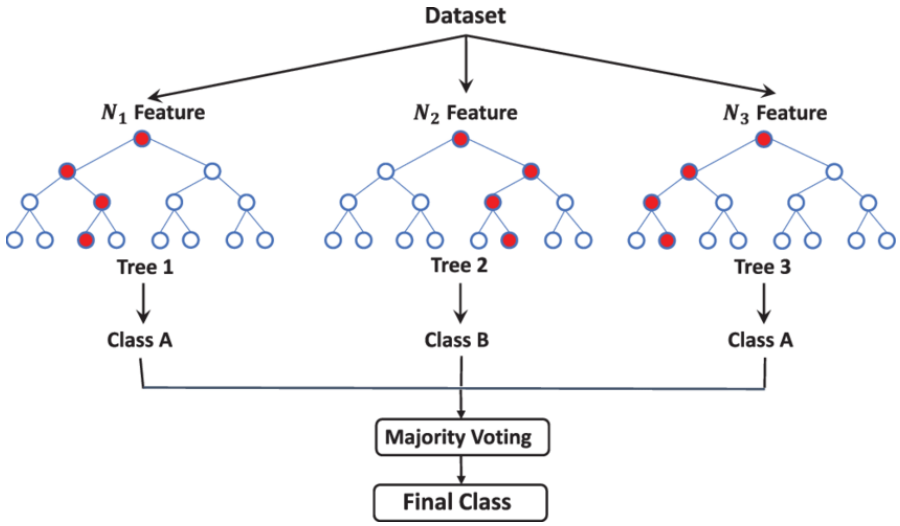

Figure 2.6: Schematic explanation of Random Forest classifier from **TAHMASEBI2020103619**. The different trees represent the decision trees constructed from the bootstrapped samples. Each tree is created on the basis of the most discriminatory features, N_x for a particular bootstrapped sample. The final class of the samples in the dataset is decided on the basis of majority voting.

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 maximally separating/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.

2.3.2 Random Forest Classifiers

Random forest Classifiers are based on the idea of bagging or bootstrap aggregation of decision trees (**hastie2009elements**). 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 speci-

fied. 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. As it is evident from the figure 2.6, the 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 it's discriminatory power. Random Forests are widely used used in Neuroimaging studies due to the interpretability of features which can be obtained using the feature ranking in terms of the feature importance.

2.3.3 Multilayer Perceptron

A Multilayer Perceptron (MLP) is an Artificial Neural Network 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 figure 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 figure 2.7 (a), the output y from the neuron can be expressed as:

$$y = g(\sum_{i=1}^{N} w_i x_i + b) \tag{2.26}$$

where w_i represents the weighting of the inputs and the 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(**universal_mlp**). Consider that we

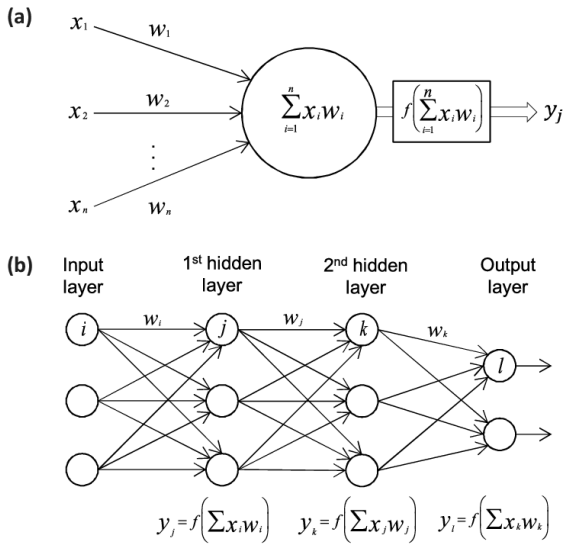

Figure 2.7: (a) Visualization of an artificial neuron from **vieira2017using** depicting the output as a function of weighted sum of the inputs (equation 2.26) (b). The artificial neural network consisting of the input layer, hidden layers and the output layer. Image from **article**

have an input vector in the N dimensional space and the output is needed in the M dimnensional space, then such an MLP (having non-linear transfer functional units) can implement any continuous mapping from the N dimensional space to the N dimensional space, with arbitrary accuracy.

The network can be made 'deep' by adding additional hidden layers. It is employed in Neuroimaging studies in use-cases where there is a requirement to eliminate the need for manual feature selection.

3 Methods

3.1 Data Acquisition

Data was acquired for 203 subjects from the s900 release of the HCP (**hcp2015wu**). Out of the total, 101 subjects were female and 102 were male. 83 females and 58 males were aged 26-30 and while the 34 male and 28 females were aged 22-25. The demographic information along with the personality trait labels were obtained from the unrestricted access data available on https://db.humanconnectome.org/

The structural and diffusion MRI files used for this project were obtained from the repositories containing volumes preprocessed using version 3 preprocessing pipelines of the HCP detailed in (**GLASSER2013105**). A Siemens 3T Skyra system was used used to scan all subjects (starting in August 2012, housed at Washington University, St. Louis). The details of the acquisition protocol are mentioned in **van2012human**.

There were two types of structural data required from the HCP pipeline for each subject in order to proceed with the task of the project. First was the segmentation volume as well as the cortical surface parcellation based on the Desikan Killinay Atlas (**desikan2006automated**, provided as a default in FreeSurfer). Second were the structural scans in undistorted native structural volume space for each subject. This was a T1w volume data in the subject's native space obtained after rigid-body rotation to AC-PC alignment (rigidly aligned to the native axis of MNI space). It was sampled at the same resolution as the diffusion data (1.25 mm isotropic, originally 0.77 mm isotropic). It was important to take into account the images from the subject's native space since it is the one in which the tractography was performed as this space is the best approximation of the subject's physical brain. The parameters of the T1w images are presented in the table 3.1.

For each subject, four types of files were used. First, the preprocessed diffusion time series file. Second, the brain mask in diffusion space. The other two types were diffusion weighting and diffusion direction for each volume. The important characteristics of the dMRI images is that they obtained in very high resolution (1.25mm isotropic) using a Stejskal-Tanner (monopolar) diffusion encoding scheme as mentioned in section 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-

TR (ms)	2400
TE (ms)	2.14
T1 (ms)	1000
Flip angle	8 deg
FOV	224x224
Voxel Size	0.77 mm isotropic

Table 3.1: Acquisition parameters for the structural image acquisition from the 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	224x224
Voxel Size	0.77 mm isotropic
b-values	100,2000 and 3000 s/mm^2

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

Parameter	Value
Sequence	Spin-echo EPI
TR	5520 ms
TE	89.5 ms
flip angle	78 deg
refocusing flip angle	160 deg
FOV	210x180 (RO x PE)
matrix	168x144 (RO x PE)
slice thickness	1.25 mm, 111 slices, 1.25 mm isotropic voxels
Multiband factor	3
Echo spacing	0.78 ms
BW	1488 Hz/Px
Phase partial Fourier	6/8
b-values	1000, 2000, and 3000 s/mm^2

Figure 3.1: Tentative, rough estimate

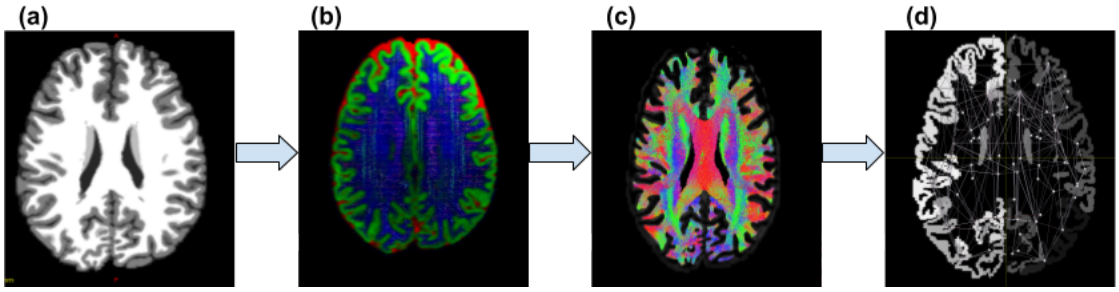

Figure 3.2: Pipeline used to create a connectome for each subject. (a) Five tissue segmented image visualized in grayscale. (b) A slice of a 4D image mapped in 3D using RGB encoding tissue densities, CSF as red, GM as green and WM as blue. (c) Fiber tractography of 1M fibers produced using probabilistic tractography overlaid on an axial slice of the brain. (d) The nodes of the connectome representing ROIs overlaid on an axial slice

value acquired once with right-to-left and another in the opposite phase encoding polarities.

3.2 Creating the Connectome

A connectome is generated on the basis of tractography as explained in section 2.1.5.The pipeline was implemented on the basis of the tutorial on Structural Connectome for Human Connectome from the software package ***Mrtrix3***. The preparation of the structural connectivity matrices can be visualized in the figure 3.2 mainly involves three steps which are explained in the following subsections.

3.2.1 Structural and Diffusion image processing

For each subject, first the structural volume sampled at the same resolution as the diffusion data (1.25mm isotropic), available in within the HCP data was selected to generate a tissue-segmented image to make it suitable for Anatomically constrained

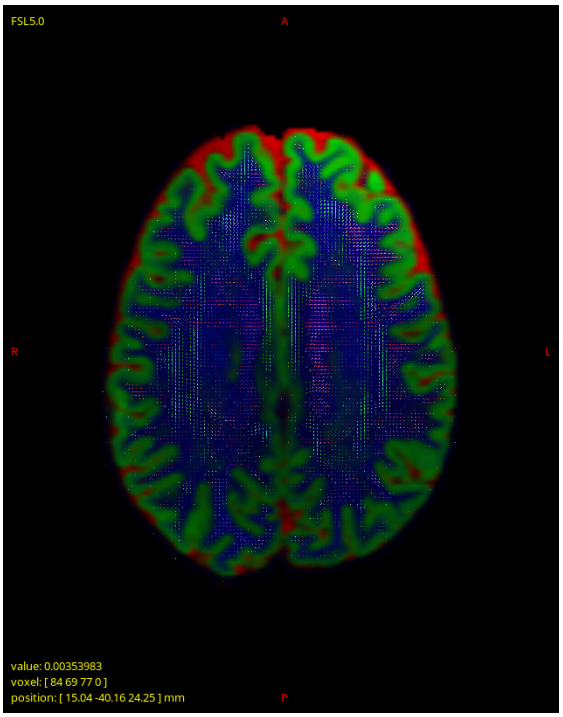

Figure 3.3: Caption

tractography (**anattractsmith**).

The diffusion image was first converted to a non-compressed format. The information about the diffusion gradient encoding was represented in the header of the file, the volume data was made continuous voxel-wise and the data points were converted to a floating point format. After this, the mean b=0 image was generated for visualization. The b=0 image serves as a sort of baseline for anatomical reference.

The multi-shell, multi-tissue response function was determined in order to form Multi-Shell, Multi-tissue spherical deconvolution. The deconvolution leads to the formation of a 4 dimensional image in which each 2D image (as viewed in 3.3) is RGB encoded where the cerebrospinal fluid (CSF) is seen in red, the gray matter in green and the white matter in blue.

3.2.2 Tractography

The initial tractography of 5M tracts was generated using ***Mrtrix3*** based on the algorithm iFOD2 (**tournier2010improved**). Anatomical constraints on the tractography were provided using the five tissue type image.The tracking along a streamline was informed to truncate and perform retracking if the fiber ter-

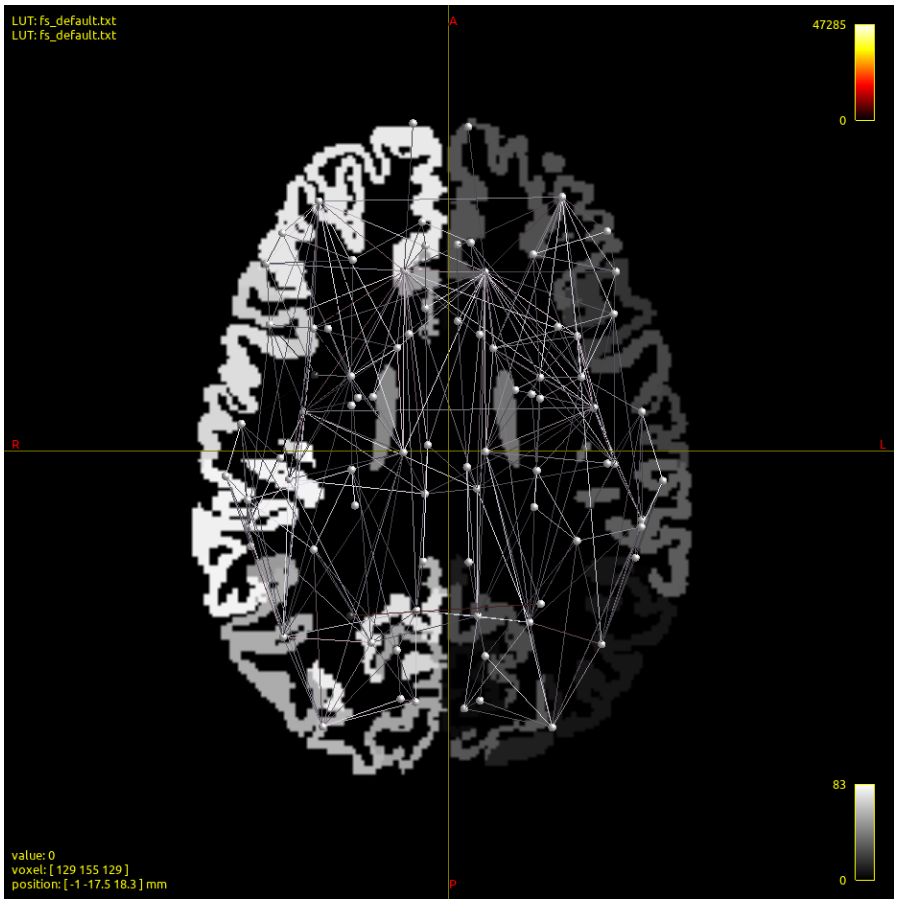

Figure 3.4: Caption

minates at poor structural termination. The streamlines were cropped whenever as the streamlines cross the grey matter-white matter interface. The seed points were dynamically determined using the Spherical-deconvolution informed filtering (SIFT= model **smith2013sift** of the white matter fODFs. The cutoff value of 0.06 was set for FOD amplitude for terminating tracks. The maximum length of streamlines was set at 250 mm (200 times the voxel size) when the voxel size is 1.25mm.

The tcksift command was used to downsample the tractography from 5M fibers to 1M fibers to preserve the most biologically relevant fibers using the SIFT algorithm (**smith2013sift**). This provides more meaningful estimates of the structural connection density.

3.2.3 Connectome Generation

The Grey matter parcellation of the scans provided in the HCP is based on the default color scheme implemented in FreeSurfer color look up table. This divides the brain into x regions which are too many for visualizing a connectome and hence they were converted to format of the lookup table defined in ***Mrtrix3***. This is the GM parcellaten obtained in as shown in Fig. . Based on structures and the parcellation schems the nodes are determined by replacing the FreeSurfer's estimates of subcortical grey matter structures with estimates from FSL's FIRST tool.

Finally, the connectome was generated for the 1M fiber tracts using on the cortical parcellation from the 84 ROIs specified by the Desikan Killiany Atlas in ***Mrtrix3***. In the Fig. it can be seen that there is a need to decipher the streamlines which are terminating in the specified regions of interest.Using different parameter settings from the tck2connectome command the features extracted were mean FA of the streamlines that connect the two regions, the mean streamline length and number of streamlines between them. The visualization can be seen in the figure 3.4. The connectome is represented in an upper triangular matrix format (.csv file) that 84x84, 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 dataframe was prepared in order to be prepared to fit into the classifier. Each row represented the data for an individual subject. While, each column represented the connection between any two brain regions i.e. features in one cell of the connectivity matrix. The exact structure varied according to the feature selection techniques employed.

The whole experiment was divided into two types of experiment. The first was in the form of a baseline where the feature selection was based on statistical scores such as fscores, t-tests and pearson correlation coefficients. This type of feature selection was termed as the 'baseline' experiments set. The second approach, added another layer to the first approach with the incorporation of MEWS problem to reduce the graphs into smaller subgraphs explained in the later section.

3.3.1 Exclusion of self loops

Self loops are considered the connections from the brain regions to themselves. On the basis of performance of the classifiers it was seen that without the self loops

the performance the classifiers still perform almost similarly.

To see the difference in between the classification accuracies, a paired sample t-test was conducted. In this t-test the paired samples are the classification metrics in the dataframe including the self loops and excluding them. The null hypothesis of this two-sided t-test was that the average of a particular classification metric on the basis of one feature (such as meanFA, mean streamline length etc.) with respect to five different personality traits was identical. The null hypothesis could not be rejected because for the test data, the p-values of these t-tests was high and hence not corresponding to any statistical significance.

3.3.1.1 Statistical Coefficients

The fscore used in the analysis is used to measure how well the particular feature distinguishes between the two classes labelled as 1 and 2. It has the expression according to the equation 2.13. The fscores served as feature filtering step for the baseline experiments as they were used by dividing the f-score distribution into percentiles and then choosing the percentage of features we want in the specified top percentile. However, for the solver based experiments the fscores for each feature was (numerically) low which made the MIP problem hard to run computationally. Even multiplying the f-scores with an order of 10^3 was not useful since the standard deviation of the f-scores was not too high (insert the standard deviation).

Similarly, an independent sample t-test was applied to consider the training subjects and the test subjects as independent samples, with the null-hypothesis that the means of the given feature for the two samples are identical i.e. $\overline{x_1} - \overline{x_2} = 0$. This feature selection was in fact quite useful for the feature selection in the baseline experiments. The p-value of the t-test could be divided into percentile distributions and the top percentiles could be chosen accordingly. However, for the solver based experiments this selection did not work well due to computational effort.

The pearson correlation coefficient in fact did work because it considers the linear nature of the personality trait coefficients. This type of feature selection is well reported in literature for Neuroimaging data considering continuous variables. It performed well for the baseline experiments as well as the solver based feature selection. For both the cases the absolute value of the pearson correlation coefficient was taken because only the correlation was important, whether it is positive or negative correlation was not a matter of concern for the analysis.

The f-score and the t-test were based on the binarization of the target variables according to the median values of the feature from the training set, this might lead

to information loss and hence their lower numerical values.

3.3.2 Feature selection

There were two types of feature selection techniques used before classification. The first represents a classical feature selection technique based on statistical coefficients while the second is based on extracting a subgraph. The Maximum edge weight subgraph techniques is based on exploiting the topological nature between the subnetworks while the first approach is based purely on numerical artefacts. To hypothesize with the connecitonist view of the brain, sugraph technique shall give rise to more interpretable results as it is based on the fact that certain brain connections are more important than others when it comes to classifying a particular type of label. The motivation to implement subgraph extraction methods is the fact that analyzing the inter-subject differences at the subnetwork level is much more easier than analyzing dense subgraph of whole brain connectivity.

3.3.2.1 Statistical analysis

Three different metrics were used to filter the features. Namely, pearson correlation coefficient, f-scores and the p-value of the t-test. The pearson correlation coefficient for the continuous values of the different personality traits, while the f-scores and t-test calculations were based on the conversion of the continuous variables into classes according to the section 3.4.1. These metrics were then used to rank the features.

The pearson correlation coefficient was computed between the values of the feature for the training subjects and their corresponding personality trait values. It was a trial to capture the linear relationship between the value of the structural brain connection and the outcome label. Using the numerical coefficients the features were selected on basis of top k% numerical values.

The fscore based selection was done by computing the fscore between thhe feature values for the training set and thresholded values of the target variables according to the medinan of the training data labels. These scores were then ranked using percentiles and fed to the classifier.

The t-test carried out was an independent two sample t-test with the null hypothesis that the training subjects come from the same population and have the same average of the numerical feature values. It was processed in a different manner as compared to the first two metrics, on the basis of the classes the training data was divided into two groups, one belonging to class 0 and the other belonging to class 1. Then the t-test was computed for each feature and the features which

gave low p-values (insert threshold) were selected.

3.3.2.2 Maximum Edge Weighted Subgraph

method:MEWS The input graphs implemented for the use-case of this thesis consisted of 84 nodes defined by the cortical parcellations based on the Desikan Killiany Atlas. The edges represent the properties of the connections between these nodes.

In this approach, there was one graph created for all subjects to get a generalized representation of connection strengths. First, from the calculation of the statistical coefficients of the brain connections (each feature) with respect to the target labels was determined. These coefficient values (one value for one connection and all subjects in the training set) formed the edge weights of the graph given as an input to the solver. Once the input graph was formed it was filtered to introduce sparsity. This sparsity was introduced using two constraint. The first being that the absolute value of the edge weights shall not be zero and that the edges will be fed into the solver if and only if the tractography of each subject contains atleast one streamline between the two nodes.

The edge weights represented the Pearson correlation coefficient (computed on the basis of the training data) of the individual feature values with the continuous personality trait variable.

The subgraph was meant to represent the most discriminative connections Based on the implementation by the authors, **DBLP:journals/corr/LobodaAS16**, the two major components are maximizing the function (**??**) and imposing constraints for the subgraph to be connected. The connectedness contraints are implemented using a Mixed Integer Programming formulation.

- $m \in [1, 84]$ specifying the number of nodes to be preserved in the output graph

Using these variables a set of equations can represent the structure of the subgraph:

$$\sum_{v=1}^{V} y_v = m \qquad \forall v \in V \tag{3.1}$$

According to equation **??**, an edge can be present in the subgraph only if both the vertices connecting the edge are included in the subgraph. Equation3.1 ensures that the specified number of nodes m are preserved.

3.4 Supervised Classification

3.4.1 Label Preparation

There were two different ways in which the continuous personality traits were converted into categorical variables.

1. All the values of a particular trait above or equal to the median value get assigned as 1.
2. Split the variable's distributions into three quantiles containing 1/3 of the data in terms of frequency.

3.4.2 Classification

- Support Vector Machines
- Random Forests

4 Results

5 Discussion

5.1 Statistical Coefficient

There were numerous possibilities for the edge weights in the input graph for the solver. It was maintained that the input graph edge weights shall represent a statistical coefficient which represents the strength of the connections between the two nodes/brain regions determined by the cortical parcellation. The coefficients chosen were the f-score, the pearson correlation coefficient and the p-value of the t-test.

The pearson correlation coefficient in fact did work because it considers the linear nature of the personality trait coefficients. This type of feature selection is well reported in literature for Neuroimaging data considering continuous variables. It performed well for the baseline experiments as well as the solver based feature selection. For both the cases the absolute value of the pearson correlation coefficient was taken because only the correlation was important, whether it is positive or negative correlation was not a matter of concern for the analysis.

The f-score and the t-test were based on the binarization of the target variables according to the median values of the feature from the training set, this might lead to information loss and hence their lower numerical values.

6 Conclusion