UTRECHT UNIVERSITY

Department of Information and Computing Science

Artificial Intelligence Master Thesis

Brain Imaging Analysis for Prediction of Developmental Problems in Pre-term Infants

First examiner:

Dr. Albert Ali Salah

Second examiner:

Dr. Ronald Poppe

External supervisors:

Dr. Maria Luisa Tataranno

Bob Waalrad

Candidate:

Shreyash Maini

In cooperation with:

UMC Utrecht

Abstract

Amplitude-electroencephalography and Magnetic Resonance Imaging are two non-invasive methods of examining neurological data. Both of these have demonstrated promising capabilities in the prediction of long-term neurodevelopmental outcomes for extremely preterm infants. Therefore, this thesis compared neurodevelopmental outcome predictions made by aEEG-EEG, MRI datasets, and their combination. Further, it tested the application of multiple feature reduction techniques to reduce model complexity. Moreover, it delved into an investigation of the affect caused by various scaling factors on the MRI dataset on the outcome predictions. Results show that with factor analysis employed on the aEEG-EEG dataset and the unscaled version of the MRI dataset, regression models achieved moderate to high performance scores (ranging from r=0.5407 to r=0.9173), while classification models achieved balanced accuracies ranging from 0.795 to 0.907. This thesis provides a basis for further research into multiple modality predictions, hinging on the ability to overcome the hurdle of data shortages.

Contents

List of Figures

List of Tables

List of Abbreviations

3D 3-Dimensional

aEEG amplitude ElectroEncephaloGram

AIC Akaike Information Criterion

BA Balanced Accuracy

BIC Bayesian Information Criterion

BISC BIelefelder SCreening

BSITD-III Bayley Scales of Infant and Toddler Development, Third Edition

CBCL Child Behaviour CheckList

CBCL-tot2y Child Behaviour CheckList total behavioural problem score

CI Confidence Interval

CognCSaNL Cognitive Composite Score

CSF CerebroSpinal Fluid

cUS cranial UltraSound

EEG ElectroEncephaloGraphy

EP Extremely Preterm

FMTRS Fine MoToR Scaled Score

FSIQ Full-Scale Intellectual Quotient

GM Gray Matter

GMTRS-NLca Gross MoToR Scaled Score

KMO Kaiser-Meyer-Olkin (Test)

KNN k-Nearest Neighbours

MR Magnetic Resonance

MRI Magnetic Resonance Imaging

MRMR Maximum Relevancy - Minimum Redundancy

MSE Mean Squared Error

NICU Neonatal Intensive Care Unit

NS No Scaling (Unscaled)

PIQ Performance Intellectual Quotient

PMA Post Menstrual Age

PS Processing Speed

SD Standard Deviation

SHAP SHapley Additive exPlanations

TEA Term Equivalent Age

TMotCScaNL Total Motor Composite Score

TVS Scaled with respect to age of infant at scan, relative to total brain volume

UMCU University Medical Center Utrecht

VIQ Verbal Intellectual Quotient

VS Scaled with respect to age of infant at scan

VVS Scaled with respect to age of infant at scan, relative to the difference of total brain volume and ventricle volume

WKZ Wilhelmina Children's Hospital

WM White Matter

WPPSI-III Wechsler Preschool & Primary Scale of Intelligence, Third Edition

Acknowledgements

1. Introduction

1.1 Background

Due to advancements in medical care in the fields of obstetrics and neonatology, the chances of survival for extremely preterm (EP) infants (i.e. born before 28 weeks of gestation) have significantly increased over the past few decades [1]–[4]. Nevertheless, those EP infants who survive face a significant risk of enduring long-lasting neurodevelopmental challenges, including cognitive and motor deficiencies [5]–[8].

The fundamental structure of the human brain is formed through a continuous developmental process that begins prior to birth and continues into adulthood. Early experiences play a crucial role in shaping this structure, laying down either a sturdy or a fragile foundation for subsequent learning, health, and behavior [9]. Infants born preterm are more susceptible to having a fragile foundation. The premature transition to an extra-uterine environment during a critical phase of rapid brain development leads to brain abnormalities in EP infants, which are the primary contributors to adverse neurodevelopmental outcomes [10]. Hence, an individualized approach to precision medicine is warranted for such fragile neonates for preventing or reducing neonatal brain injury. As a result, there is a growing focus on identifying accurate and sensitive brain-based markers that can predict the future outcomes of EP infants while they are still in the neonatal intensive care unit (NICU) [11]. By doing so, it becomes feasible to provide precision medicine which includes personalized care and timely interventions to safeguard the vulnerable developing brains of EP infants, ultimately reducing long-term complications in this population.

This thesis builds upon the work of Wang et al. and expands their framework to include an MRI (Magnetic Resonance Imaging) dataset along with

aEEG (amplitude ElectroEncephaloGram) of the cohort [12]. Furthermore, it explore the possibilities of using more powerful feature reduction techniques in an effort to make the predictions of neurodevelopmental outcomes more accurate and robust.

1.2 Research Motivation

To understand the evolving landscape of neonatal neurodevelopment, there arises a necessity to investigate the possible advantages associated with aEEG-EEG. As neonates mature, an expanded array of clinical, neuromonitoring, and neuroimaging data, including MRI scans, becomes accessible. The integration of these diverse modalities presents an opportunity to refine predictive models.

The body of literature examining the predictive role of aEEG in assessing outcomes for preterm infants continues to grow exponentially. Nevertheless, several important issues need to be addressed to enhance our understanding of this domain.

There is a lack of studies specifically targeting extremely preterm infants [13]. Since aEEG characteristics evolve weekly during the preterm period, findings from other age groups cannot be directly applied to EP infants [14].

The timing of neurodevelopmental evaluations varies among different studies, and internationally standardized scales are not consistently employed. This inconsistency complicates the comparison and generalizability of findings across studies. Existing research, moreover, typically extracts aEEG characteristics during a later postnatal period, several weeks after birth, which may not be ideal for implementing early interventions like stem cell therapy [15]. Therefore, it is necessary to explore the predictive value of aEEG features obtained during the first few days after birth.

Even when EP infants are included, however, the sample sizes are often too small to draw definitive conclusions, and variations in procedures and analyses among studies make it challenging to compare results [16].

When it comes to MRI, recently there has been a major focus into the research regarding factors that can lead to a range of neurodevelopmental disorders. To explore the relationship between the brain and behavioral changes in neonates, noninvasive neuroimaging methods such as MRI are crucial. Due to the relatively smaller size of cerebral structures and the incomplete maturity of tissues in neonates, it is critical to use age-specific MR (Magnetic Resonance) sequencing. This is necessary as the signal and contrast characteristics in neonatal brains differ significantly from those in adult brains [17].

Overall, addressing these issues will contribute to advancing our knowledge regarding the role of aEEG and MRI in predicting outcomes for preterm infants.

1.3 Project Characteristics

1.3.1 Research Question and Objectives

The objective of this project is to create a machine learning pipeline for the prediction of nuerodevelopmental outcomes for preterm infants.
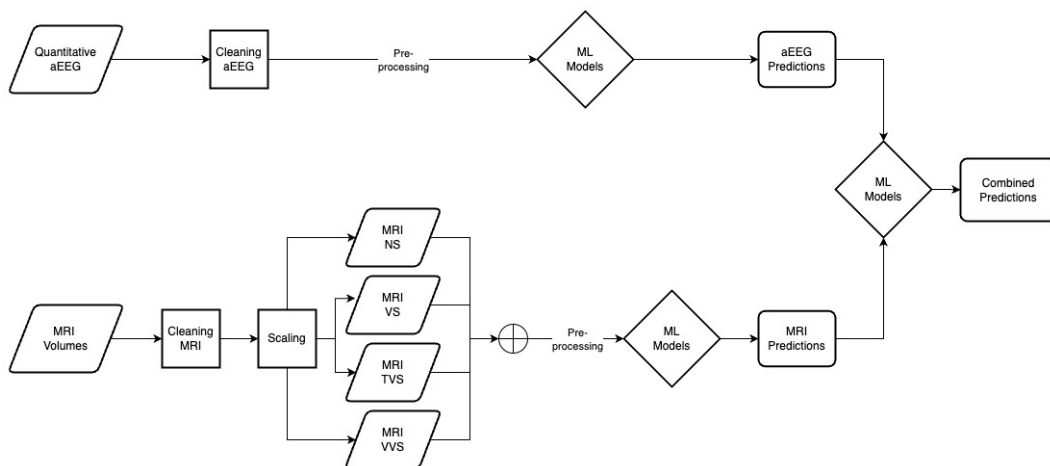
Figure 1.1: Research steps for developing the prediction models

The Figure 1.1 shows both the datasets undergo a cleaning process. The aEEG dataset goes through a pre-processing step that includes outlier detection, imputation, and normalization. It is then passed forward to either a

regression or classification Machine Learning model to get predictions. The MRI dataset is scaled to form 4 different instances of the dataset. One of these instances is then passed through the same type of pre-processing as aEEG. It is used to train either a regression or classification Machine Learning model to get predictions. Both the predictions are again fed to a regression or classification model to generate compounded predictions.

The above pipeline is implemented to answer the following questions:

1. Can employing better feature reduction techniques reduce the complexity of the predictions of neurodevelopmental outcomes in preterm infants?

2. Is it possible to improve the prediction of neurodevelopmental outcomes in neonatal preterm infants by involving MRI data?

1.3.2 Scope

The project scope is defined as designing the prediction pipeline using machine learning algorithms that provide the highest standard of interpretability. To achieve the best possible predictions, a search is conducted within the features present in the aEEG and MRI datasets. Finally, a comparative analysis is conducted to evaluate if a combination of the two datasets gives better predictions than either dataset alone.

2. Theoretical Background

A foundation of the knowledge and a literature review is provided in this chapter. Firstly, the anatomy of a neonatal infant brain is explained, followed by modelling methods of both neonatal aEEG and MRI as well as their suitability for outcome prediction, and their accompanying limitations.

2.1 Neonatal Cranial Anatomy

The head of a human infant comprises of a skull, CerebroSpinal Fluid (CSF) and the brain.

The cranium (or skull) is a composite structure made of bone, which surrounds and protects the brain. The CSF is a clear, colourless body fluid found in the brain and spine that acts like a buffer to help cushion the brain and spinal cord from injury, removes waste products, and serves to regulate intracranial pressure.
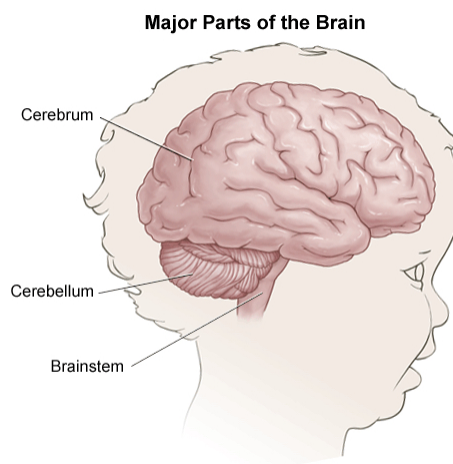
Figure 2.1: Parts of the brain (Taken from Anatomy of a Child's Brain - Stanford Medicine Children's Health)

The brain is the focal organ of the human nervous system. It is composed

of the cerebrum, cerebellum, and brainstem (Figure 2.1). The cerebellum's function is to coordinate muscle movements, maintain posture, and balance. The brainstem acts as a relay center connecting the cerebrum and cerebellum to the spinal cord. It also performs many autonomous functions such as body temperature and heart rate regulation. The cerebrum is the largest part of the brain and is composed of right and left hemispheres. These hemispheres have distinct fissures, which divide the brain into 4 lobes: frontal, temporal, parietal, and occipital. Each of these can be further divided into areas that serve very specific functions [18].

2.2 Modelling the Neonatal Brain

2.2.1 aEEG in Neonatology

Continuous electroencephalography (EEG) monitoring using a full array of electrodes is considered the most reliable method for monitoring brain activity [19]. EEG data can be mathematically represented as a time series of voltage measurements recorded from multiple electrodes placed on the scalp. Each electrode corresponds to a specific channel, and the voltage values captured at regular intervals over time form the EEG signal. Mathematically, EEG data can be denoted as a sequence of voltage values:

$$EEG(t) = [V_1(t), V_2(t), V_3(t), ..., V_n(t)] \tag{2.1}$$

where EEG(t) represents the EEG signal at time t, $V_1(t)$ to $V_n(t)$ represent the voltage measurements at each electrode/channel at that specific time point. EEG signals are typically sampled at a specific sampling rate, denoted as f_s (measured in samples per second). The sampling rate determines how many voltage measurements are recorded per second and affects the temporal resolution of the EEG signal.

Continuous EEG monitoring is highly valuable for tracking ongoing brain activity and detecting seizures in newborns.

Amplitude-integrated electroencephalogram (aEEG) is a simplified and non-invasive method of monitoring cerebral function at the bedside using one or two channels [20]. Its ease of use and effectiveness has led to increasing utilization of aEEG, along with access to the corresponding raw EEG traces, in neonatal intensive care units (NICUs) to assess brain function in preterm infants [20], [21]. The aEEG can be initiated soon after admission to the NICU, enabling early detection and intervention for potential brain dysfunction [22]. Significantly, both qualitative and quantitative parameters of the aEEG have demonstrated promising capabilities in predicting long-term outcomes for preterm infants [13], [15], [20], [23]–[28].

aEEG exhibits prime performance in predicting neurodevelopmental outcomes [29] and has the ability to identify preterm infants who may be at risk of poor outcomes even without obvious brain injury [15].

aEEG is used in intensive care settings to assess brain function over extended durations, often spanning hours to days, surpassing the temporal scope of traditional electroencephalogram (EEG) recordings. aEEG signals, derived from a modified version of the EEG, can similarly be measured by the traces of electrical activity detected by electrodes placed on the scalp of a patient. aEEG can be represented as a compressed display of the EEG signal over time in the form of a trace or a graph. The vertical axis represents the amplitude or voltage range, while the horizontal axis represents time.
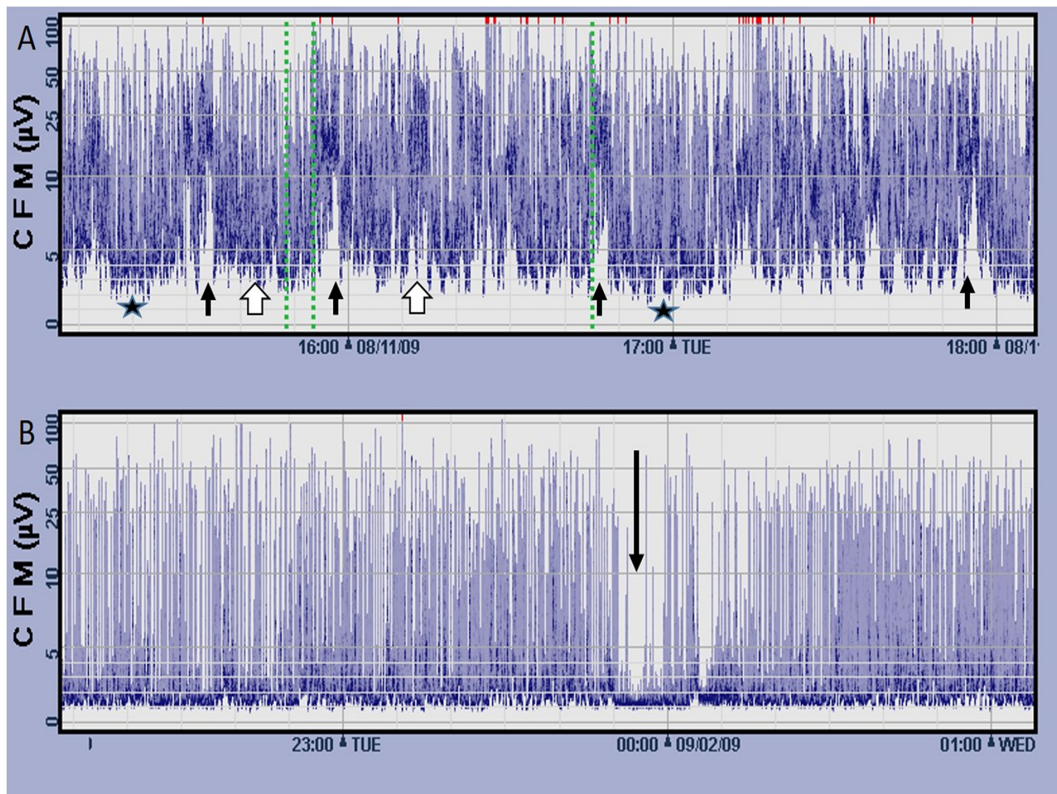
Figure 2.2: Example of aEEG recordings from an extremely preterm infant [30]

In contrast to the full EEG signal, aEEG typically involves a reduced number of channels or electrodes (e.g., one or two channels) to simplify monitoring and analysis.

Typically, electrodes that measure EEG are arranged according to the international 10-20 system [31]. This method establishes consistent testing procedures, enabling the compilation, replication, and effective analysis and comparison of study outcomes using the scientific method. It relies on the correlation between the electrode placement and the corresponding region of the cerebral cortex, ensuring an accurate and standardized assessment of brain activity. "10-20" refers to the percentages of skull length and width that are distances between the electrodes. As can be seen in Figure 2.3, each site of electrode placement corresponds to a letter to identify the area of the brain it reads. These include: pre-frontal (Fp), frontal (F), temporal (T), parietal (P), occipital (O), and central (C). They are coupled with odd numbers to refer to the left side of the head, even numbers for the right side of the head or the letter "z" for zero which corresponds to an electrode placed on

the midline of the skull.
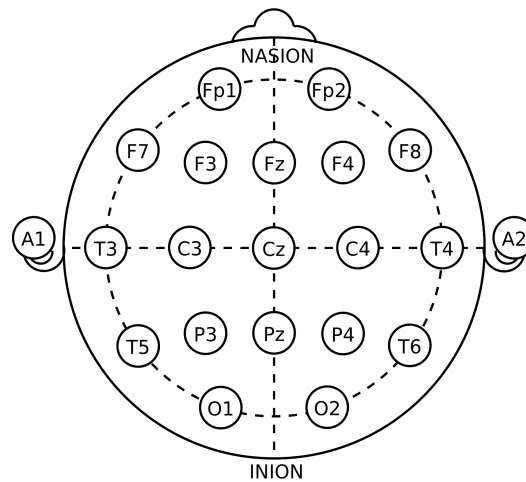
Figure 2.3: Electrode locations of International 10-20 system for EEG recording

For neonates, aEEG measures either one or two channels with the application of fewer electrodes compared to the standard for adults.

2.2.1.1 Prediction of Neurodevelopmental Outcomes with EEG

The paper by Lloyd et al. poses the question of the viability of EEG in predicting neurodevelopmental outcomes for preterm infants [28]. They captured the EEG data at three specific time points throughout the neonatal period for infants with a gestational age of less than 32 weeks. The EEGs underwent a visual interpretation assessment using a standardized grading system, which involves evaluating temporal organization/cyclicity, normal features, abnormal waves, and abnormal features to assign a grading. The grading system was then bifurcated into two categories, distinguishing between normal and abnormal EEG patterns. This dichotomous classification system was then compared with the outcome measures to draw the main conclusions of the study. After controlling for potential confounding factors, the EEG grading system demonstrated statistically significant results in predicting neurodevelopmental outcomes. It is important to note that while the grading system was objective, it still relied on interpretation by expert reviewers. This limitation is commonly acknowledged in EEG studies. Nonetheless, this particular assessment scheme stands out from previous grading systems due to the inclusion of detailed definitions in the user in-

structions manual. This inclusion significantly reduces subjectivity among reviewers, enhancing the reliability of their results.

Another study by Burger et al. followed 248 preterm infants and assessed their aEEG for dominating background activity, calculation of the percentage of continuous activity for each period and infant, qualitative visual assessment Burdjalov score and the minimum, mean and maximum amplitude parameter values [23]. These values were then compared to the Bielefelder screening (BISC) performances of the group at five years of age. It was shown that the minimum amplitude and the percentage of continuous background activity was considerably lower at postnatal week two and three, respectively, in infants with poor BISC performance. These findings show significant evidence for the claim that aEEG in preterm infants was able to predict neurodevelopmental outcomes as well as precursor skills of literacy at later stages in life.

The study by Wang et al. investigates whether aEEG features obtained within the first three days after birth could predict neurodevelopmental outcomes at the age of two and five, in a large and homogeneous cohort of extremely preterm (EP) infants [12]. They extracted a range of qualitative and quantitative aEEG features for the analysis of outcome prediction. Initially, employing machine learning-based regression models to compare findings with previous studies that used group-level statistics such as multiple regression analysis. Additionally, they utilized Machine Learning-based classification models to differentiate between EP infants with delayed or optimal outcomes, aiming to provide findings that were more clinically applicable. The goal was to explore effective prognostic tools for long-term disability and aid in delivering timely and personalized interventions in the neonatal intensive care unit (NICU) setting.

2.2.2 MRI in Neonatology

MRI or Magnetic Resonance Imaging is a noninvasive neuroimaging technology that produces three dimensional detailed anatomical images. In neonatology, it used to establish links between the brain and behavioral

changes in newborns and infants. MRI is powerful, non-ionising and has several advancements to evaluate preterm brains. It is considered the best method to detect white matter injuries and enables early diagnosis of lesions [32], [33].
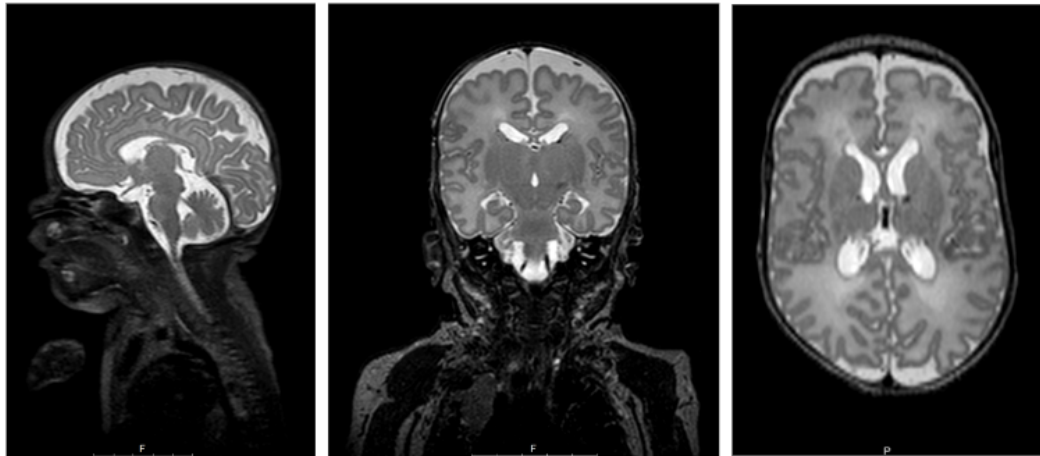
Figure 2.4: Sagittal, Coronal and Axial/Transversal MRIs

Although, MRI does have some caveats attached to it. It is a more expensive, time-consuming and challenging method, and sources conflict over their accessibility in NICUs [34], [35]. Certainly, though, neonates do need to be moved to a separate scanner from the NICU. This can be dangerous for unstable infants like preterm neonates. This arises the question: why not utilize cranial ultrasound (cUS) instead? It is inexpensive, similarly non-invasive, and can be performed by a variety of trained professionals [35]. cUS is superior than MRI in the detection of acute intraventricular haemorrhage, perforator stroke and cerebral sinovenous thrombosis. It is noteworthy, however, that the majority of infants experiencing neurodevelopmental deficits do not display these typical injury patterns. This renders the sensitivity of cUS (cranial UltraSound) in predicting outcomes poor. MRI steps in and surpasses cUS in its capability to detect subtle brain abnormalities [36]. Moreover, MRI facilitates comprehensive measurements of brain regions' volumes and growth rates, including but not limited to the white matter, ventricular system, cortex, deep gray matter, and cerebellum. These regions, frequently altered following preterm birth, are more thoroughly examined through MRI techniques [37].

2.2.2.1 Prediction of Neurodevelopmental Outcomes with MRI

Recent studies emphasize the vital role of MRI in detecting brain abnormalities at 2 years of age, especially in preterm infants. Woodward et al. studied the connections between white-matter(WM) and gray-matter(GM) abnormalities on MRI and the risks of severe impairment [38]. Moderate-to-severe cerebral WM abnormalities present in infants were predictive of adverse outcomes on cognitive and motor functions, and of cerebral palsy.

Previously, the cerebral WM was primarily looked at as the site of injury [39], [40]. Newer research, however, highlights the involvement of GM and the cerebellum as well [41], [42].

Using a 3D volumetric approach to MRI, Kidokoro et al. evaluated MRI scans of very preterm infants, scoring the severity of brain injuries and impaired growth in the WM, GM and cerebellum [43]. They concluded that WM signal abnormalities were associated with a reduced deep gray matter area but not with cerebellar abnormality. In addition, that hemorrhages were associated with cerebellar signal abnormality and volume reduction. The scoring system they devised provides a comprehensive and objective classification of brain abnormalities.

Inder et al. review three major forms of brain injury in very preterm infants that can lead to adverse neurodevelopmental outcomes [44]. They acknowledge the modified brain development in preterm infants and yield a comprehension of pivotal factors influencing the period spent in the NICU. They further add that adverse neurodevelopmental consequences in preterm infants stem not only from brain injury but also from impaired brain development, occurring independently of injury. They outline the possible areas of injury and dysmaturation that correlate with the outcomes, as well as underline potential interventions to improve them.

Advanced MRI techniques not only reveal signal abnormalities but also define overall and specific changes in brain structure in children who were born preterm. Therefore, a comprehensive evaluation utilizing MRI is essential to understand the complete impact of preterm birth on the developing brain, encompassing neurodevelopmental deformities.

3. Methodology

Figure 3.1: An infant in the NICU (© 2020 Eddie Lawrence/Science Photo Library)

3.1 Data Acquisition

The data for this thesis had been previously collected by the Wilhelmina Children's Hospital of the University Medical Center Utrecht as a part of the 'Preparing for RSV Immunisation and Surveillance in Europe (PROMISE)' research project.

This retrospective cohort study included extremely preterm infants (Gestational age < 28 weeks + 0 days) who were admitted to the NICU of the Wilhelmina Children's Hospital (WKZ) in Utrecht, The Netherlands, between October 2006 and September 2018. The infants included in the study had undergone continuous two-channel aEEG monitoring within the first three days after birth. Additionally, between 36 and 42 weeks of postmenstrual age (PMA), the babies also underwent MR Imaging. Infants with genetic or metabolic diseases, and major congenital malformations, were excluded from the study. The study obtained permission from the Medical Research Ethics Committee (abbreviation in Dutch: METC) of the University Medi-

cal Center Utrecht (UMCU) to use patient data. Since the aEEG data used were collected as part of routine medical care and the analysis was retrospective, written parental consent was not required. For the collection of the MRI data, oral informed parental consent was obtained. All data were pseudonymized and de-identified prior to any analysis. An overview of the data is given in Table 3.1.

	Extremely preterm infants (n=380)
Maternal and demographic characteristics	
Sex	
Female	171 (45%)
Male	209 (55%)
Gestational age at birth, in weeks	26.43 (1.1, 23.86-27.86)
Birth-weight, in grams	875.5 (180.5, 460-1480)
Maternal education level	
No education	1 (<1%)
Primary education	12 (3%)
Some secondary education	36 (9%)
Completed secondary education	84 (22%)
University education	101 (27%)
Missing	146 (38%)
Clinical characteristics during NICU stay	
Morphine	
Yes	220 (58%)
No	146 (38%)
Missing	14 (4%)
Cumulative dosage, in mg/kg	1.3 (2.3, 0-18.6)
The administration of anti-seizure, sedative, or anaesthetic medications	
Yes	154 (41%)
No	226 (59)
Illness severity	
Severe	182 (48%)
Mild	185 (49%)
Missing	13 (3%)
Presence of severe brain injury	
Yes	116 (31%)
No	236 (62%)
Missing	28 (7%)
Apgar score	
At 1st min after birth	4.9 (2.3, 0-9)
At 5th min after birth	7.1 (1.7, 1-10)
At 10th min after birth	8.2 (1.1, 3-10)
Follow-up age	
Age at the time of BSITD-III administration, in years	2.2 (0.25, 1.9-2.7)
Age at the time of CBCL administration, in years	2.2 (0.25, 1.9-2.7)
Age at the time of WPPSI-III administration, in years	5.8 (0.2, 5.4-6)

Table 3.1: Patient characteristics

In the table 3.1, data is presented as "count (%)" for qualitative, and as

"mean (Standard Deviation (SD), range)" for quantitative columns. The anti-seizure, sedative, or anaesthetic medications include phenobarbital, lidocaine, levetiracetam, clonazepam, midazolam, and other potential surgical anaesthetics. Illness severity is defined based on the total number of days of mechanical ventilation during the NICU stay: severe for >=7 days, mild for <7 days. The following number of infants were missing: Birth-weight: 1; Morphine dosage: 14; Apgar at 1st: 8; Apgar at 5th: 5; Apgar at 10th: 87.

3.1.1 aEEG-EEG

As part of the standard care provided to extremely preterm (EP) infants in the NICU at WKZ, bedside two-channel aEEG monitoring was initiated promptly after admission and continued for the first three days following birth. The BrainZ monitor (BRM2 or BRM3; Natus Medical Inc., Seattle, WA) was utilized for the aEEG recording. Raw EEG signals were recorded from needle electrodes placed subcutaneously in pairs (i.e. channels) over the left (F3-P3) and right (F4-P4) frontoparietal cortex. The sampling rate was set at 256 Hz, and these signals were subsequently processed to generate aEEG traces. Further, a reference electrode was positioned over the vertex (Cz). The placement of electrodes was performed by nurses or clinicians in accordance with the International 10-20 system.
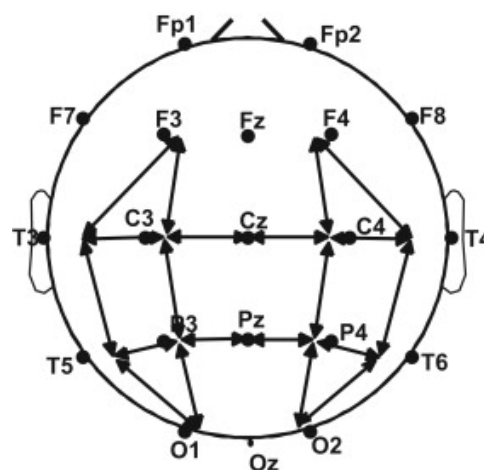
Figure 3.2: Location and nomenclature of the electrodes on an infant's head

In the course of this research, the dataset containing aEEG information

underwent a comprehensive processing pipeline, detailed in the study led by Wang et al. The initial aEEG was passed through an algorithm that identified and eliminated any artifacts present in its raw signals. Furthermore, the pipeline also involved feature extraction. Through this, meaningful and relevant information was distilled from the raw aEEG signals. This dataset comprises of 330 distinct features (110 for each day of processed aEEG) from 369 patients.

3.1.2 MRI

MR images were acquired around term-equivalent age (TEA) on a 3.0 Tesla MR system (Philips Healthcare, Best, The Netherlands) using a sense head coil [45]. TEA refers to the age at which a prematurely born infant's brain development matches that of a full-term infant. Since premature infants are born before their brains have fully developed, their MRI scans may be assessed and compared to those of term infants at a certain age to gauge their developmental progress. Studies have demonstrated that TEA-MRI correlate with neurodevelopmental outcomes, offering enhanced predictive accuracy compared to other measures such as neuroimaging, clinical assessments, or physical examinations [34].

The dataset for MRI includes volumes for all the different parts of the brain along with their volumes relative to the total volume of the brain. The 277 columns contain the patient ID, age of the patient at scan, the total volume of the brain, volumes of 120 different parts of the brain, the relative volumes of the same 120 parts of the brain and 34 columns from the Kidokoro system of analysis. This dataset comprises records from 264 patients.

3.1.3 Target Variables

The neurodevelopmental progress of EP infants was assessed during their routine follow-up appointment at the outpatient clinic of the WKZ when they were between 2 and 3, and again between 5 and 7 years old. The kids were rated by medical professionals who had no knowledge of their prior aEEG-MRI characteristics. The assessments were made using Dutch coun-

terparts of the following:

- *Bayley Scales of Infant and Toddler Development, Third Edition (BSITD-III)* was administered to the preschool age toddlers to evaluate cognitive and motor functioning. Four index scores were provided by the BSITD-III—cognitive composite score, total motor composite score, and fine and gross motor sub-scaled scores.

- *Child Behavior Checklist (CBCL)* was used to assess behavioural or emotional problems, yielding a total problem score.

- *Wechsler Preschool & Primary Scale of Intelligence, Third Edition (WPPSI-III)* was used to assess cognitive abilities of early school age children by providing a full-scale intellectual quotient (FSIQ), and composite scores for verbal IQ (VIQ), performance IQ (PIQ), and processing speed (PS).

Lower scores on the BSITD-III, and WPPSI-III indicate poorer functioning, whereas higher scores on the CBCL indicate more behavioural or emotional problems. For classification, each outcome was converted into a binary form (optimal vs impaired) using the thresholds in Table 3.2.

	Optimal	Impaired
BSITD-III	>10 %ile	<=10 %ile
CBCL	<70	>=70
WPPSI-III	>70	<=70

Table 3.2: Thresholds for optimal vs. impaired classification

This system of classifying gives the data distribution given in Fig 3.3.
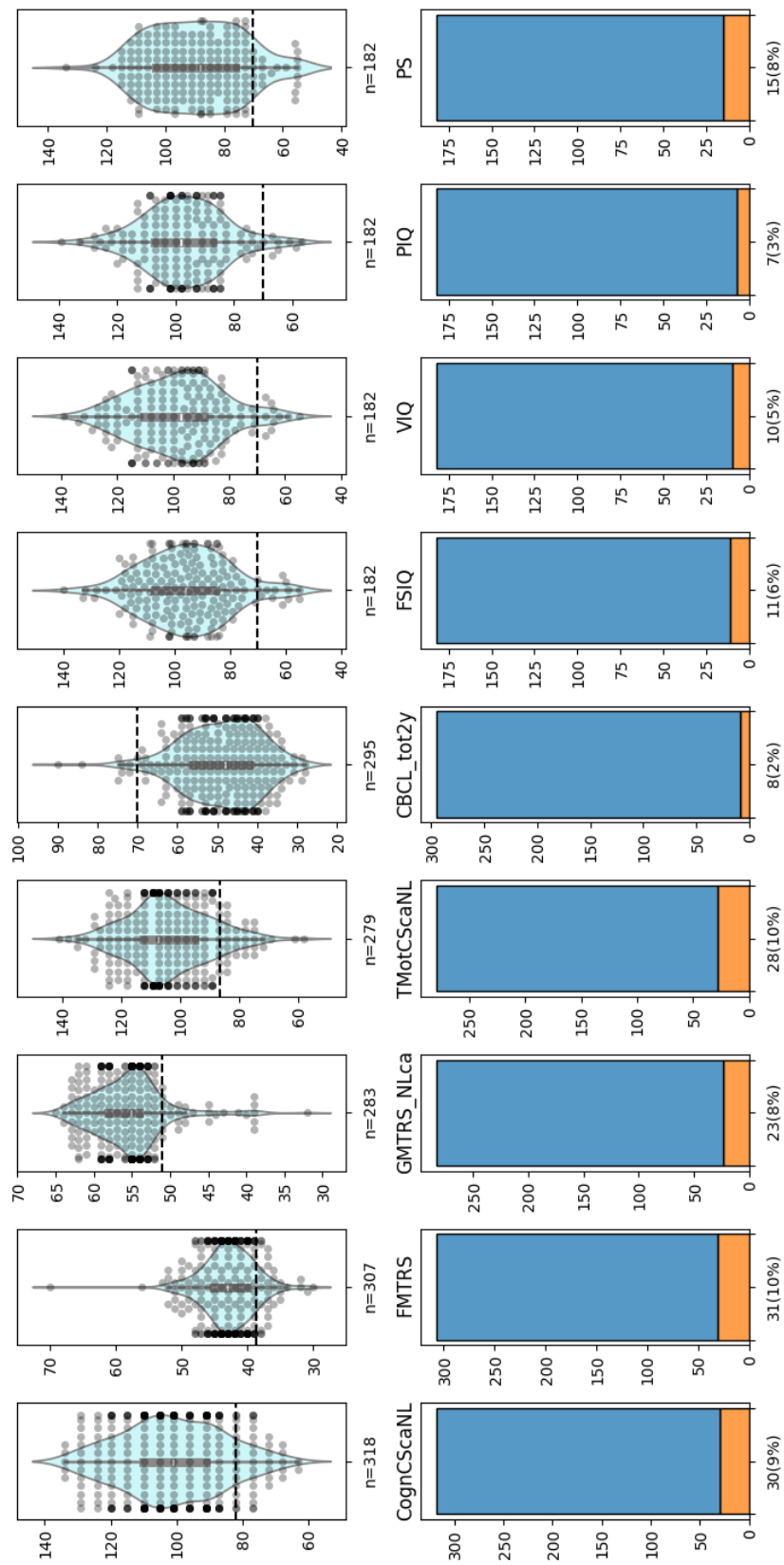
Figure 3.3: Distribution of neurodevelopmental outcomes

On each violin plot, the central gray line denotes the mean, the vertical thick line represents the SD, and the horizontal dash line indicates the threshold value that divides the optimal and impaired classes. The number of infants who had outcome measurements is indicated below each violin plot. The number (and percentage) of infants classified as impaired is given below each bar plot.

3.2 Exploratory Data Analysis

Two spearman's correlation coefficient matrices are given in figure 3.4. 3.4 (a) shows a 339 × 339 matrix illustrating the pairwise relationships between the 330 EEG features and 9 outcomes, while, 3.4 (b) shows a 283 × 283 matrix with 274 MRI features. Within the matrices, the intensity of the colour gradient corresponds to the strength of the correlation, with lighter shades indicating positive correlations and darker shades indicating negative correlations. The correlations between individual features of either dataset and outcomes were relatively weak.

(a)aEEG-EEG   **(b)**MRI
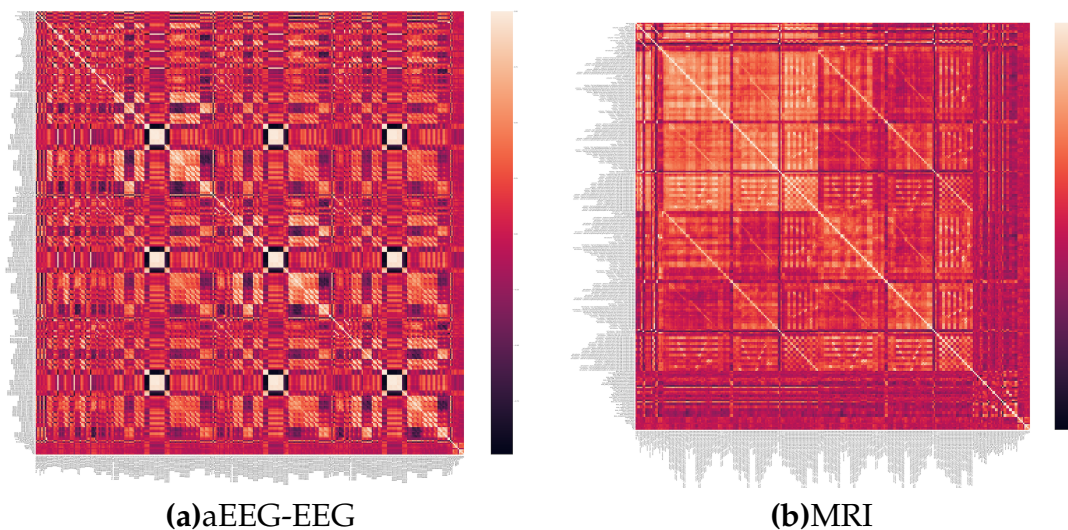
Figure 3.4: Correlation Matrices

A cohort of 601 extremely preterm infants underwent screening. Among these, 108 were excluded due to either absence of data or presence of pre-congenital birth defects. 113 more were excluded due to invalid or poor

quality of data. Subsequently, 380 infants met the eligibility criteria for the study on aEEG-EEG. The demographics and clinical characteristics of these infants during their NICU stay is presented in Table 3.1. Out of these, 11 infants did not have any aEEG or MRI data associated with them and hence were also removed from further analysis.

Of the 369 infants with eligible data, 342 (93%) had available aEEG–EEG data for all three days (i.e., $20 - 24$ hours, 44˘48 hours, and 68˘72 hours), whereas the remaining possessed data for one (2 [1%]) or two (22 [6%]) of the specified days. Notably, each infant yielded a total of 339 aEEG-EEG features comprising nine qualitative and 330 quantitative features. In this thesis, only quantitative features are analysed.[1]

From the total 601 preterm infant cohort, 493 were imaged using MRI. Of these, only 264 (54%) were found to be eligible for this study. The rest were excluded due either invalid or substantial lack of data.
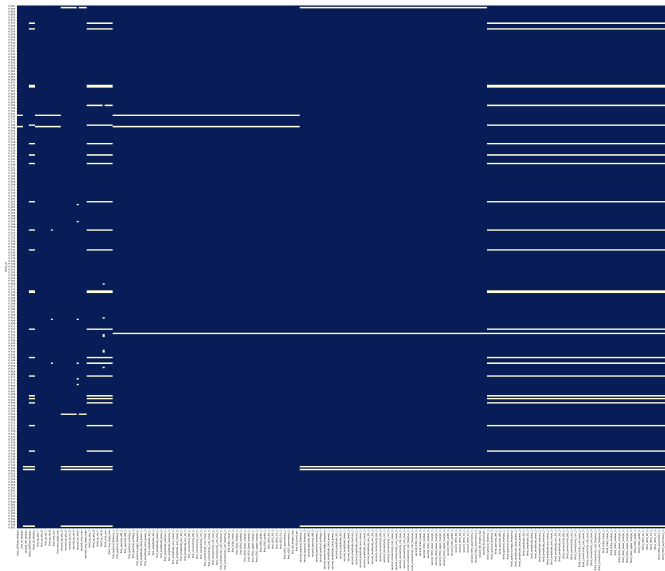
Figure 3.5: Null values in the aEEG-EEG dataset

---

[1]The distributions of the quantitative features across the three periods are delineated in Appendix A.
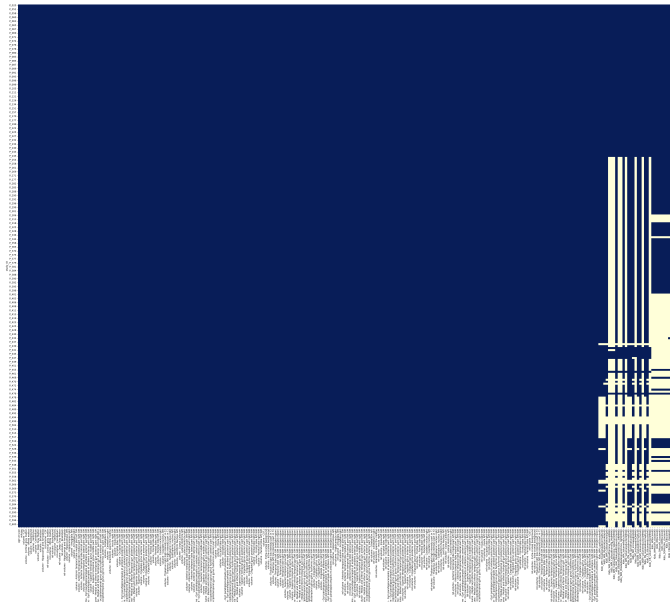
Figure 3.6: Null values in the MRI dataset

Figures 3.5 and 3.6 show null values present in the datasets as yellow and non-null values as blue.

Initially, as shown in figure 3.3, the following number of outcomes have been measured:

- BSITD-III: Maximum of 318 (cognitive) and a minimum of 279 (total motor) with an average of 10% impairment

- CBCL: 295 scores with 2% impairment

- WPPSI-III: 182 scores with an average of 6% impairment

When accounting for the data in the datasets, however, there is outcome data available that corresponds to patient IDs that have been excluded. An outline of outcome data matching patient IDs currently in the analysis is represented in Table 3.3 for aEEG data and in Table 3.4 for MRI data.

Outcome	Measured outcomes	Impaired outcomes	Impaired %
Cognition	271	25	9
Fine motor	261	29	11
Gross motor	237	20	8
Total motor	233	26	11
CBCL	247	7	3
Full scale IQ	151	10	7
Verbal IQ	151	10	7
Performance IQ	151	7	5
Processing speed	151	14	9

Table 3.3: Number of outcomes available for aEEG-EEG analysis

Outcome	Measured outcomes	Impaired outcomes	Impaired %
Cognition	199	21	10
Fine motor	192	21	11
Gross motor	171	16	9
Total motor	168	19	11
CBCL	189	5	3
Full scale IQ	98	7	7
Verbal IQ	98	7	7
Performance IQ	98	5	5
Processing speed	98	8	8

Table 3.4: Number of outcomes available for MRI analysis

3.3 Data Preprocessing

This section documents all the steps followed to transform and structure the data into a format that could be utilized in the machine learning algorithms that model the data subsequently.

3.3.1 Cleaning Data

Cleaning the data refers to the process of identifying and correcting errors and inaccuracies in an effort to ensure the integrity of the dataset for analysis. This step streamlines the integration and further analyses of the MRI and aEEG datasets by standardizing column names. This standardization also facilitates merging these datasets with another that includes target variables.

3.3.1.1 aEEG-EEG

A part of this preprocessing step is removing unnecessary variables from the datasets to focus on relevant ones. From the aEEG dataset, in particular, it excludes any remark columns. Although the remarks provided by the doctors who administered the patients in the study may contain valuable insights for future medical professionals managing the patient's care, they are deemed unsuitable for inclusion in the current research scope. The exclusion is justified by the subjective nature of these remarks, as they might represent qualitative observations rather than quantifiable statistics.

3.3.1.2 MRI

For MRI, two datasets are provided. One including Kidokoro scores and one including the volumes and age at scan of infants. All the volumes present in the second dataset are standardized as numeric data. This involves the handling various formats and decimal point systems, allowing for consistent treatment of numerical values. The two datasets are then merged to form a single MRI superset that includes the age of the infants, volumes of all the parts of the brain and the Kidokoro scores involved with them.

Furthermore, the dataset includes columns for relative volumes, i.e., unaltered volumes relative to the total brain volume. These exhibit high multicollinearity. Multicollinearity occurs when two or more independent variables in a model are highly correlated, making it challenging to distinguish their individual effects on the dependent variable. Consequently, the relative volumes will be removed during this step. Further down the pipeline, more collinear variables will also be removed.

3.3.2 Scaling Data

The MRI dataset includes the volumes of different parts of the brain, along with the ages of the infants on the day of the scan. As neonates grow, their brains undergo rapid developmental changes, and accounting for these changes can be essential for accurate analysis.

Allometric scaling is a mathematical approach that can evaluate the non-

uniform relationship between observations or quantities (e.g., brain volumes) when comparing across different species. It involves adjusting a scaling exponent to better capture the proportional changes in the variable of interest with respect to size or age.

According to Wang et al., different allometric exponents for drug dosages were identified depending on the age range [46]. The lowest similarity in the exponents was within neonatal populations. Their findings show that for scaling clearance, while a singular exponent for allometric scaling may be of value for scaling from adults to paediatric patients, the same cannot be said for scaling within neonates.

In the context of brain volumes in neonates, different age groups exhibit differing scaling exponents that would accurately represent brain growth patterns. The total brain volume, hence, becomes an essential factor for determining the scaling. Using this, the volumes of the various parts of the brain can be scaled proportionately to the age of an infant.

For this scaling, the following formula is used:

$$V_{scaled} = V_{original} + m * (40 - PMA) \tag{3.1}$$

Here, V_{scaled} is the scaled value of the volume of a part of the brain, $V_{original}$ is the original value of the volume of a part of the brain, 40 is the average age in weeks for a non-preterm infant, PMA refers to the Post Menstrual Age, i.e., the infant's age in weeks when the scan occurred, and finally, m is the slope of the volume of the part of the brain. Furthermore, when calculating the slope of the volumes, a 1-D α trimming filter was applied. This refers to the exclusion of the highest and lowest $\alpha\%$ from the analysis in an attempt to mitigate the influence of outliers on the slope.
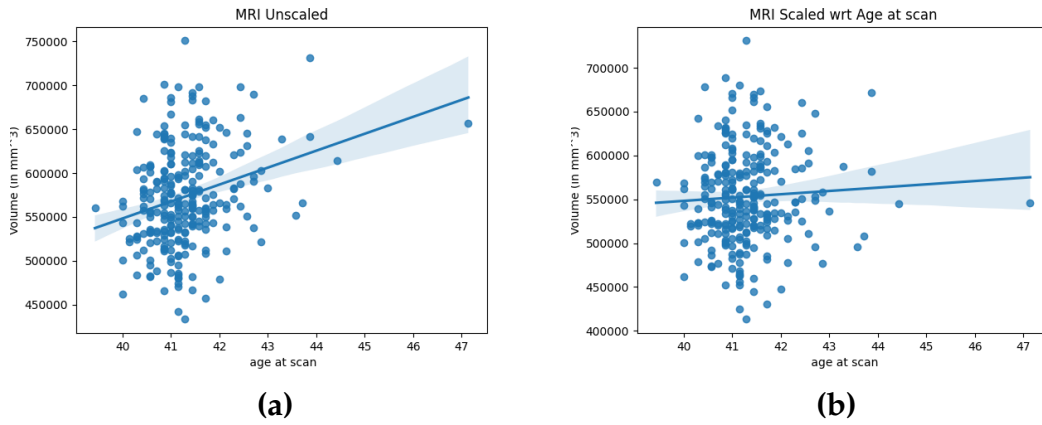
Figure 3.7: Total volume of the brain vs. age of infant at scan **(a)** before and **(b)** after scaling

The volumes are then divided to get volumes relative to the total brain volume, and the total brain volume - ventricle volume. In summary, the following variations of the MRI dataset will be compared further:

1. V_{NS} - No scaling (NS)

2. V_{VS} - Scaling w.r.t. age of infant at scan (VS)

3. V_{TVS} - Scaling w.r.t. age, relative to total brain volume (TVS)

4. V_{VVS} - Scaling w.r.t. age, relative to the difference of total brain volume and ventricle volume (VVS)

3.3.3   Outlier Detection

Outlier detection involves identifying and potentially removing data points that deviate significantly from the average of a dataset. For each data point in a dataset, we compute its z-score, Z, which measures how many standard deviations it is away from the mean of the dataset. The formula for calculating the z-score of a data point, X, in a dataset with mean μ and standard deviation σ, is:

$$Z = \frac{X - \mu}{\sigma} \tag{3.2}$$

If a data point has a z-score close to 0, it means it is close to the mean. Data points with z-scores outside a certain threshold range (chosen as, ± 3) are flagged as potential outliers. These flagged data points are then imputed.

3.3.4   Data Imputations

Missing data can occur due to various reasons such as human error during data entry, equipment malfunction, or even the nature of the data collection process. Data imputation is a technique used to fill in these missing or incomplete values within a dataset. The following methods of imputation were considered:

1. **Zero-Value Imputation**

   This involves imputing the value of the column to 0. The main reason why this can be preferred is to avoid the introduction of potential biases. It ensures minimal distortion to the original dataset.

2. **Minimum Value Imputation**

   Imputing missing values with the minimum observed value in a dataset. It helps maintain the original range of values. The limitations of imputing missing values with the minimum is that it might underestimate true values and misrepresent the nature of the data.

3. **Mean Value Imputation**

   Imputing missing values with the mean of the observed data helps maintain the central tendency of the dataset and keeps the overall average unchanged. Imputing missing values with the mean comes with the disadvantage that it might underestimate the variability of the data.

4. **Median Value Imputation**

   Imputing missing values with the median helps maintain the shape and characteristics of the distribution. The median is less sensitive to outliers compared to the mean. Median imputation, however, does not utilize all available information in the dataset, potentially leading to a loss of variability and information.

5. **Mode Value Imputation**

   Imputing missing values with the mode helps maintain the original data characteristics by filling in the missing information with the most prevalent data point. Mode imputation is less suitable for continuous or numeric variables where it is challenging to define a mode, as continuous variables often have unique or continuous distributions.

6. **Nearest Neighbours Imputation**

   KNN (k-Nearest Neighbours) imputation estimates missing values based on similarities or distances between data points. It uses the values of nearby or similar instances to impute the missing values, leveraging existing relationships within the dataset. KNN imputation is versatile and applicable to both continuous and categorical data. It is imperative, however, to choose K judiciously and consider computational constraints when applying KNN imputation.

3.3.4.1   Checking validity

To check the validity of the various imputation methods, a k-fold cross validated linear regression was performed on the data. This involves splitting the dataset into k subsets called 'folds'. The model is then trained k times, each time on a separate fold and evaluated on the remaining folds. The process provides multiple evaluations of the same dataset. The results are averaged to get more reliable performance metrics. The following regression metrics were used to ascertain which imputation method would be the most suitable:

The r^2 score is a measure of the proportion of variance in the target variable that is explained by the independent variables used to train a model. This value can range from -infinity to 1, where:

- 1 indicates that the model is a perfect fit explaining all of the variance.

- 0 indicates that the model does not explain any variance

- <0 indicates that the model performs worse than a simple mean-based model.

The MSE or Mean Squared Error is the average of the squared distances between the actual and predicted values of the target variable.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad (3.3)$$

Where n is the number of samples, y_i is the actual value, $\hat{y}_i$ is the predicted value. MSE penalizes larger errors more heavily than smaller errors. The lower the MSE, the better the performance of the model.

This validity check for imputations was conducted on unadulterated aEEG and MRI datasets. It was found that the median imputation had the best performance across all the runs. Therefore, in any further analysis, median imputations were utilized.

3.3.5 Normalization

Similar to the technique applied in outlier removal, the data is normalized using z-score. Also known as standardization, it is a statistical technique used to transform data into a standard normal distribution, particularly useful when dealing with features that have vastly differing scales and units. Z-score normalization scales the data so that it has a mean of 0 and a standard deviation of 1.
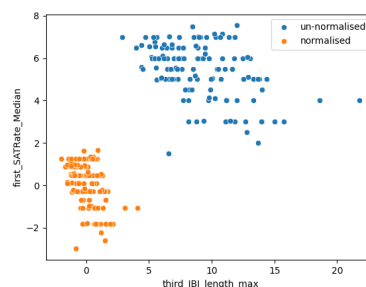
Figure 3.8: An example of z-normalization

3.3.6 Feature Reduction

3.3.6.1 Factor Analysis

Before we utilize factor analysis, the suitability of the data must be assessed. This is usually done via the tests described below.

The Kaiser-Meyer-Olkin (KMO) Test gives a measure of sampling adequacy that indicates the proportion of variance in variables that might be caused by underlying factors. The higher the proportion, the higher the KMO-value, the more suited the data is to factor analysis. Its value is given by:

$$KMO_j = \frac{\sum_{i \neq j} R_{ij}^2}{\sum_{i \neq j} R_{ij}^2 + \sum_{i \neq j} U_{ij}^2} \tag{3.4}$$

where, R_{ij} is the correlation matrix and U_{ij} is the partial covariance matrix. KMO values vary from 0 to 1. The KMO values between 0.8 to 1.0 indicate the sampling is adequate, between 0.7 to 0.79 are middling and values between 0.6 to 0.69 are mediocre. KMO values less than 0.6 indicate the sampling is not adequate and remedial action should be taken [47]. This remedial action refers to removing these features from our dataset as they are too correlated and would not offer any more substantial information to the analysis.

The second test conducted is the Bartlett's test of Sphericity. It is used to test the null hypothesis that the variables are orthogonal. If the null hypothesis is true, that would mean that the correlation matrix is an identity matrix. If so, the variables tested are unrelated and not ideal for factor analysis. If there are k samples with sizes n_i and sample variances S_i^2 then Bartlett's test statistic is given by:

$$\chi^2 = \frac{(N - k) \ln(S_2^p) - \sum_{i=1}^{k} (n_i - 1) \ln(S_i^2)}{1 + \frac{1}{3(k-1)} \left(\sum_{i=1}^{k} \left(\frac{1}{n_i - 1} \right) - \frac{1}{N - k} \right)} \tag{3.5}$$

where, $N = \sum_{i=1}^{k} n_i$ and $S_2^p = \frac{1}{N-k} \sum_i (n_i - 1) S_i^2$ is the pooled estimate for the variance [48].

If the data has passed the above two tests, it can be concluded that a feature reduction technique like factor analysis would be able to capture significant variance present in the data within a condensed number of variables.

With the remaining features, we need to find out the number of factors they need to be reduced to. One method to do this is to look at the cumulative explained variance graph. Usually, the number of factors that are enough to explain 99% of the total variance in the dataset are chosen.

Factor analysis is a statistical method used to explain the variability of correlated features. It can be used to draw a lower number unobserved variables called factors, that can describe the same variation as a higher number of observed variables. Factor analysis searches for linear combinations of variables that can form factors. The factor loading of a variable is the quantification of the extent to which the variable is related to a given factor. [49]

The technique attempts to explain a set of p observations in each of n individuals with a set of k common factors ($f_{i,j}$) where there are fewer factors per unit than observations per unit ($k < p$). Each individual has k of their own common factors, and these are related to the observations via the factor loading matrix $L \in \mathbb{R}^{p \times k}$, for a single observation, according to

$$x_{i,m} - \mu_i = l_{i,1} f_{1,m} + \cdots + l_{i,k} f_{k,m} + \varepsilon_{i,m} \tag{3.6}$$

where, $x_{i,m}$ is the value of the ith observation of the mth individual, μ_i is the observation mean for the ith observation, $l_{i,j}$ is the loading for the ith observation of the jth factor, $f_{j,m}$ is the value of the jth factor of the mth individual, and $\varepsilon_{i,m}$ is the $(i, m)^{th}$ unobserved stochastic error term with mean zero and finite variance.

When the maximally loaded features are assigned to factors, certain factors are left empty. These factors are "empty" as no further combination of

features explains enough variance in the data.

The factor loadings give the weighting of each feature in the linear equation that forms a factor. These factors can then be used as the new features that are used to train subsequent models.

3.3.6.2 MRMR

Maximum Relevancy - Minimum Redundancy (MRMR) is a feature selection algorithm first proposed in the paper by Ding et al. and expanded upon by Zhao et al. for Uber [50], [51]. MRMR is a minimal-optimal feature selection algorithm. This means that it seeks the smallest set of features that have maximum predictive possibilities. Working iteratively, MRMR identifies the features that exhibit the highest relevance to the target variable while minimizing redundancy with previously selected features.

The Uber paper describes multiple variants of MRMR, however, it concludes FCQ (F-test correlation quotient) to be one of the best. It is highly valuable not only due to its effectiveness but also because of its simplicity, which makes it a fast and easily implementable technique in any downstream machine learning model.

f is the importance of a feature X_i ($i \in \{1, 2, \cdots, m\}$), given by:

$$f^{FCQ}(X_i) = \frac{F(Y, X_i)}{\frac{1}{|S|} \sum_{X_s \in S} \rho(X_s, X_i)} \tag{3.7}$$

The numerator is the relevancy of a feature X_i, computed as the F-statistic, F, between the target variable Y and the current feature X_i. The denominator is the redundancy, computed as ρ, the Pearson correlation, between the set of all previously selected features X_s and the current feature X_i, averaged over the length of selected feature set S.

3.4 Regression

Linear regression is a statistical method that fits a linear equation to the relationship between a dependent variable, y, and one or more independent variables, x.

$$\hat{y} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \varepsilon \tag{3.8}$$

The equation above is for a multiple linear regression, where, β_0 is the intercept, β_i is the slope for x_i, which is the ith independent variable, and ε is the error term, i.e. the difference between the predicted value $\hat{y}$ and the observed value y of the dependent variable.

Machine learning based linear regression is utilized to model the relationship between the dependent variable and one or more independent variables. The term "linear" refers to the assumption that any change in x is associated with a linear change in y. What gives it a basis in Machine Learning, is the training phase. In it, the algorithm "learns" from the training dataset and adjusts the parameters of the linear equation to minimize the error between the actual and predicted values.

3.5 Classification

Logistic regression is a statistical method used to predict the probability of a categorical outcome based on one or more independent variables. Despite being called "regression", it is actually a classification algorithm. To model the probability of a binary outcome, it fits a sigmoid function to a linear combination of independent variables.

$$P(y = 1|x) = \frac{1}{1 + e^{-\hat{y}}} \tag{3.9}$$

The equation states the the probability P of the dependent variable y belongs to the class 1, given the independent variables x is the sigmoid function applied to the linear model $\hat{y}$. The logistic regression fits the linear model coefficients β to minimize the residual sum of squares between the observed and predicted target variables. Once the model is trained, it can be used to predict the probability of an instance belonging to a certain class. If this probability is greater than a chosen threshold, it is classed as y belonging to 1, else it will be classed as y belonging to 0.

3.6 Explanation of models

To explain the machine learning models applied in this thesis, SHAP (SHapley Additive exPlanations) is utilized.

The Shapley value is a solution concept in cooperative game theory, providing a possible answer to the question of how important each player is to the overall cooperation of a team. In machine learning, they provide a principled way to explain the predictions of models. The calculated value is the average marginal contribution of a feature value across all possible coalitions. This value is calculated by:

$$\varphi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(n - |S| - 1)!}{n!} (v(S \cup \{i\}) - v(S)) \qquad (3.10)$$

where S is a coalition of features, N is the set of n total number of features and the sum extends over all subsets S of N not containing feature i, including the empty set. $v(S)$, called the worth of coalition S describes the total expected sum of payoffs the members of S can obtain by cooperation.

SHAP assigns each feature used to train the model this importance value φ for its predictions, providing a way to compute which features contribute the most to a prediction.

4. Results

4.1 Feature Reduction

4.1.1 Factor Analysis

Initially, the KMO test is conducted. For the aEEG dataset, out of the 330 total features, 144 passed with a KMO value of over 0.6. These obtained an average score of 0.9 among the accepted variables. For MRI, out of the 154 total features, the number of features passing with a KMO value of 0.6 for NS was 60, for VS was 60, for TVS was 56, and for VVS was 55. These accepted fields obtained an average score of 0.887 for NS, 0.879 for VS, 0.783 for TVS, and 0.774 for VVS. For the sake of consistency, the maximum number of passed features (i.e. 60) are delivered to the next step. In addition to the datasets that did not have 60 columns, certain features as determined by experts at the University Medical Center Utrecht (UMCU) were added. Total brain volume, ventricle volume and global Kidokoro score are namely some of the features that were appended.

Secondly, Bartlett's test of Sphericity is conducted. For the aEEG dataset, the χ^2 value is 98547.854 and a p-value that rejects the null hypothesis. While for the MRI dataset, the χ^2 value is 29481.973 for NS, 28485.747 for VS, 25574.715 for TVS, and 24258.364 for VVS, and p-values that reject the null hypothesis for each.

It is important to note that the two tests were conducted on a subset of variables in the dataset. This was done to mitigate the impact of multicollinearity which would result in a singular matrix with a determinant of 0, eventually leading to divide by zero errors.
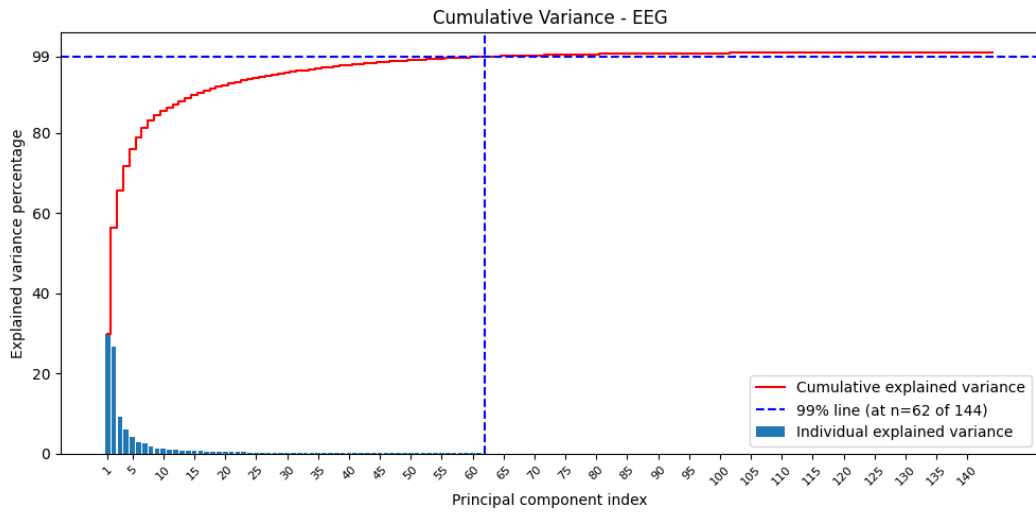
(a)
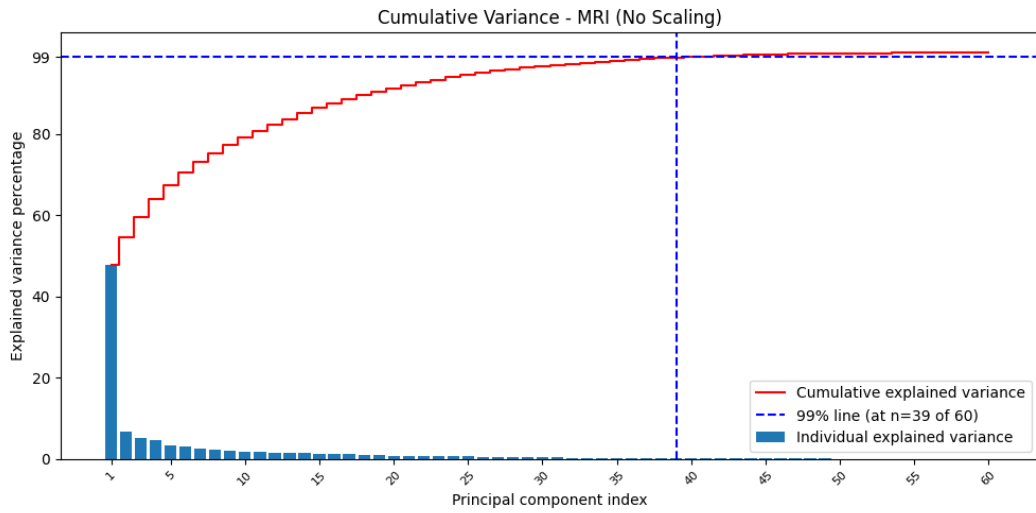
(b)
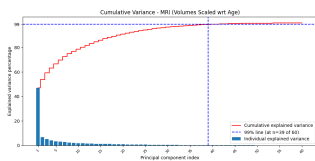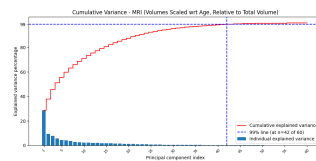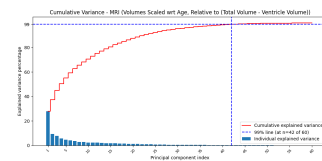
(c)   **(d)**   **(e)**

Figure 4.1: Cumulative variance graphs for each of the tested datasets

Thirdly, the cumulative explained variance graph is drawn. As can be seen from Fig 4.1, the red line plots the cumulative explained variance, and the point where the two dotted lines meet is at 99% of cumulative variance and the number of factors that should be chosen. For the aEEG dataset, this

point is at 62 and for MRI: NS is 39; VS is 39; TVS is 42; VVS is 42. To keep number of features consistent for MRI comparisons, 42 will be chosen.

Finally, the features are factor analysed and the loadings are deduced as shown in figure 4.2. In **(a)** and **(b)** of the figure, the features are sorted on the basis of the factors they maximally load on to. Whereas, in **(c)**, **(d)** and **(e)**, the features remain sorted as they were in **(b)** to show the differences that different scalings can make in factoring the MRI dataset.
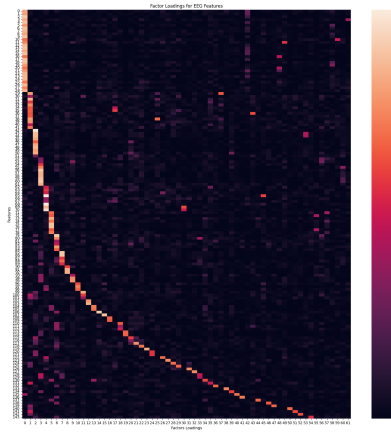
(a)aEEG-EEG
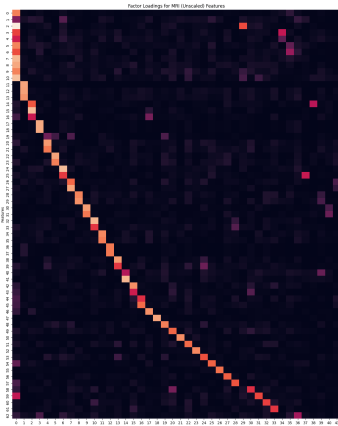
(b)MRI (Unscaled)
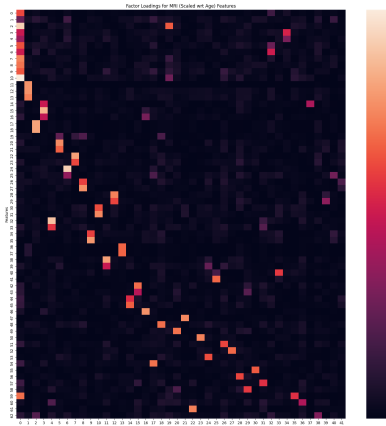
(c)MRI (Scaled)
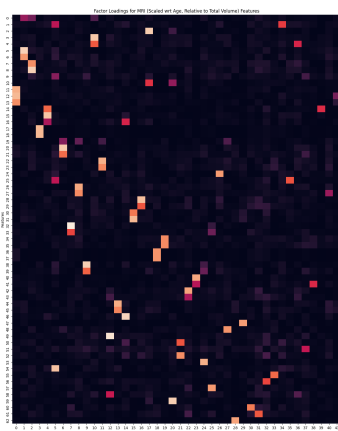
(d)MRI (Scaled, Rel. total volume)
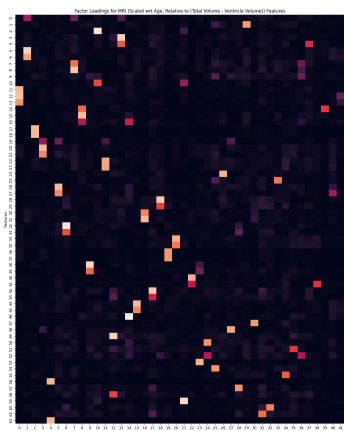
(e)MRI (Scaled, Rel. total-ventricle volume)

Figure 4.2: Factor loadings for each of the tested datasets

Assigning maximally loaded features to factors, it is found that the aEEG dataset has 45 explainable factors, while MRI has: NS: 35; VS: 36; TVS: 38; VVS: 38 explainable factors.

4.1.2 MRMR

As concluded from the third step of Factor Analysis, 62 features from aEEG and 42 features from each MRI dataset are chosen using the MRMR selection algorithm.[1]

4.2 Regression Results

A linear regression model was trained on each variation of the datasets. To compare against the various models, metrics from a leave-one-out-cross-validation evaluation were used. Mean Squared Error was used as the measure to calculate the AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) for each model.

---

[1] The list of features selected by MRMR is given in Appendix E.

Target Variable	Feature Reduction	aEEG-EEG		MRI NS	
		AIC	BIC	AIC	BIC
CognCScaNL	None	645.75	1838.051	195.75	706.213
	Factor Analysis	80.54	246.238	60.751	179.31
	MRMR	114.632	341.566	74.683	216.295
FMTRS	None	652.257	1832.113	221.131	726.043
	Factor Analysis	85.816	233.703	66.718	175.34
	MRMR	119.875	344.44	80.892	220.964
GMTRS_NLca	None	653.65	1801.578	204.584	691.542
	Factor Analysis	85.136	229.112	64.69	163.82
	MRMR	119.299	337.786	79.39	214.481
TMotCScaNL	None	645.79	1788.084	251.071	735.286
	Factor Analysis	80.858	222.887	60.836	157.393
	MRMR	114.966	332.381	75.288	209.619
CBCL_tot2y	None	646.299	1807.907	206.303	708.773
	Factor Analysis	82.11	228.213	62.602	168.502
	MRMR	115.93	337.022	76.44	215.835
FSIQ	None	647.749	1646.469	295.824	696.494
	Factor Analysis	79.882	196.874	59.72	139.089
	MRMR	113.62	303.709	73.515	184.668
VIQ	None	647.997	1646.716	295.882	696.552
	Factor Analysis	79.568	196.559	59.06	138.428
	MRMR	113.639	303.727	73.895	185.049
PIQ	None	648.052	1646.772	295.472	696.142
	Factor Analysis	80.613	197.604	60.706	140.075
	MRMR	114.112	304.2	73.607	184.761
PS	None	647.948	1646.667	295.097	695.767
	Factor Analysis	79.441	196.432	59.257	138.626
	MRMR	113.763	303.851	73.543	184.697

Table 4.1: AIC/BIC scores for linear regression (Best Models)

The model with the lowest values for both turned out to be Factor Analysis across the board. With aEEG-EEG scoring an average value of 81.552 AIC and 216.402 BIC, and MRI scoring 61.593 AIC and 155.62 BIC.

To gauge the statistical significance of the aforementioned models, permutation tests were employed. In these tests, the scores of the outcomes were subject to random shuffling iteratively, a total of 1000 times per test. The metrics tested were the Mean Squared Error (Eq. 3.3), and the Pearson correlation metric r.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \qquad (4.1)$$

The value of r ranges from -1 to 1, where 1 denotes a perfect match, 0 denotes a chance prediction, and negative values indicate progressively poorer performance.

4.2.1   aEEG-EEG

aEEG-EEG	r	p (r)	mse	p (mse)
BSITD-III				
Cognitive composite	0.3746	<0.001	211.144	<0.001
Fine motor scaled	0.5755	<0.001	12.158	<0.001
Gross motor scaled	0.5235	<0.001	15.5018	<0.001
Total motor composite	0.5539	<0.001	132.4329	<0.001
CBCL				
Total behavioural problem	0.4775	<0.001	72.1681	<0.001
WPPSI-III				
Full-scale IQ	0.7888	<0.001	90.0927	<0.001
Verbal IQ	0.7895	<0.001	92.3642	<0.001
Performance IQ	0.7854	<0.001	74.7971	<0.001
Processing speed	0.6788	<0.001	132.7127	<0.001

Table 4.2: Prediction performance of linear regression on aEEG-EEG

The prediction performance of the significant models was moderate for BSITD-III (r=0.3746 − 0.5755) and CBCL (r=0.4775), and high for WPPSI-III (r=0.6788 − 0.7895). Conversely, the composite scores for cognition, total motor, and processing speed did have high mean squared errors (mse=211.144, 132.4329, 132.7127).

(a) Regression plot for CognCScaNL

(b) Regression plot for PS

(c) SHAP values for CognCScaNL

(d) SHAP values for PS
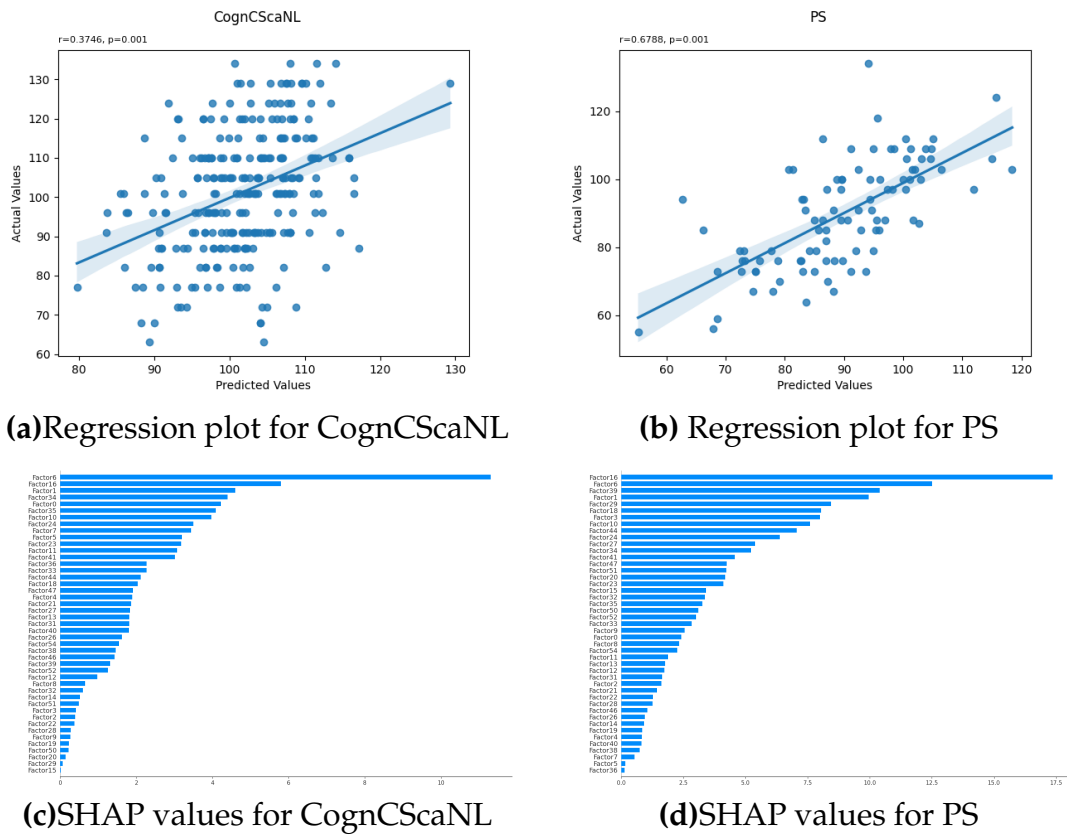
Figure 4.3: Regression results for aEEG-EEG for cognitive composite score and processing speed score

In Figure 4.3 (a) & (b), the x-axis shows the predicted values, the y-axis shows the actual values, the solid line indicates linear regression fit, and the shaded area indicates a 95% confidence interval.

Examining the top 5 SHAP values for the factors most influential to the results, the following is observed[2]:

- *Cognition composite score:* Factor 6, Factor 16, Factor 1, Factor 34 and Factor 0, correspond to a linear fusion of features with statistics of aEEG range (day 1), range (day 3), range (day 2), amplitude (day 3), and connectivity (day 2/3), respectively.

- *Processing speed score:* Factor 16, Factor 6, Factor 39, Factor 1 and Factor 29, correspond to a linear fusion of features with statistics of aEEG range (day 3), range (day 1), spectral content (day 1), range (day 2),

---

[2]All features corresponding to the factor are present in Appendix D.

and spectral content (day 2), respectively.

4.2.2 MRI

MRI NS	r	p (r)	mse	p (mse)
BSITD-III				
Cognitive composite	0.4883	<0.001	195.5259	
Fine motor scaled	0.5189	<0.001	8.0684	<0.001
Gross motor scaled	0.628	<0.001	17.7953	<0.001
Total motor composite	0.6066	<0.001	128.4436	<0.001
CBCL				
Total behavioural problem	0.5218	<0.001	65.1569	<0.001
WPPSI-III				
Full-scale IQ	0.7986	<0.001	78.2548	<0.001
Verbal IQ	0.7649	<0.001	105.6402	<0.001
Performance IQ	0.8199	<0.001	52.4974	<0.001
Processing speed	0.6634	<0.001	146.3863	<0.001

Table 4.3: Prediction performance of linear regression on MRI (Unscaled)

The prediction performance of the models on the MRI dataset was better than aEEG-EEG for all BSITD-III (r=0.4883 − 0.628, mse=8.0694 − 195.5259), CBCL (r=0.5218, mse=65.1569) and WPPSI-III (r=0.7649 − 0.8199, mse=52.4974 − 146.3863).

(a) Regression plot for CognCScaNL

(b) Regression plot for PS

(c)SHAP values for CognCScaNL

(d)SHAP values for PS
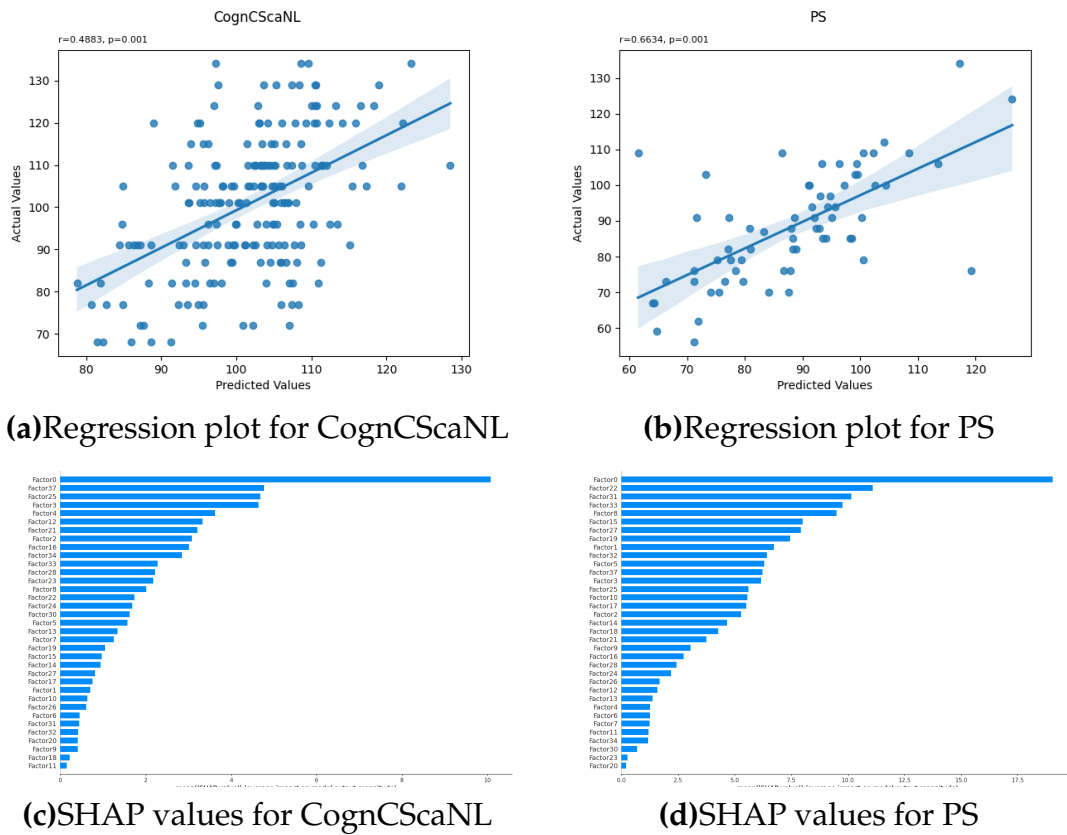
Figure 4.4: Regression results for MRI for cognitive composite score and processing speed score

In Figure 4.4 (a) & (b), the x-axis shows the predicted values, the y-axis shows the actual values, the solid line indicates linear regression fit, and the shaded area indicates a 95% confidence interval.

Examining the top 5 SHAP values for the factors most influential to the results, the following is observed[3]:

- *Cognition composite score:* Factor 0, Factor 37, Factor 25, Factor 3 and Factor 4, correspond to a linear fusion of volumes of GM, entire brain, Anterior temporal lobe GM, CSF, and thalamus (high intensity), respectively.

- *Processing speed score:* Factor 0, Factor 22, Factor 31, Factor 33 and Factor 8, correspond to a linear fusion of volumes of GM, WM, Superior temporal gyrus GM, Cingulate gyrus WM, and Gyri parahippocam-

---

[3]All features corresponding to the factor are present in Appendix D.

palis WM, respectively.

4.2.3 Combination

Combination	r	p (r)	mse	p (mse)
BSITD-III				
Cognitive composite	0.5976	<0.001	156.4118	<0.001
Fine motor scaled	0.6424	<0.001	7.3358	<0.001
Gross motor scaled	0.7417	<0.001	12.5239	<0.001
Total motor composite	0.755	<0.001	74.1583	<0.001
CBCL				
Total behavioural problem	0.5407	<0.001	54.3845	<0.001
WPPSI-III				
Full-scale IQ	0.8923	<0.001	35.2361	<0.001
Verbal IQ	0.8508	<0.001	47.077	<0.001
Performance IQ	0.9013	<0.001	28.928	<0.001
Processing speed	0.9173	<0.001	46.3915	<0.001

Table 4.4: Prediction performance of linear regression on the combination of aEEG-EEG and MRI (Scaled wrt age, relative to (total volume - ventricle volume))

The predictions of the models of the combination of aEEG and MRI predictions outperformed both aEEG and MRI alone for all the outcome predictors. BSITD-III: r=$0.5976 - 0.755$, mse=$7.3358 - 156.4118$; CBCL: r=0.5407, mse=54.3845; WPPSI-III: r=$0.8508 - 0.9173$, mse=$28.928 - 46.3915$.

(a) Regression plot for CognCScaNL

(b) Regression plot for PS
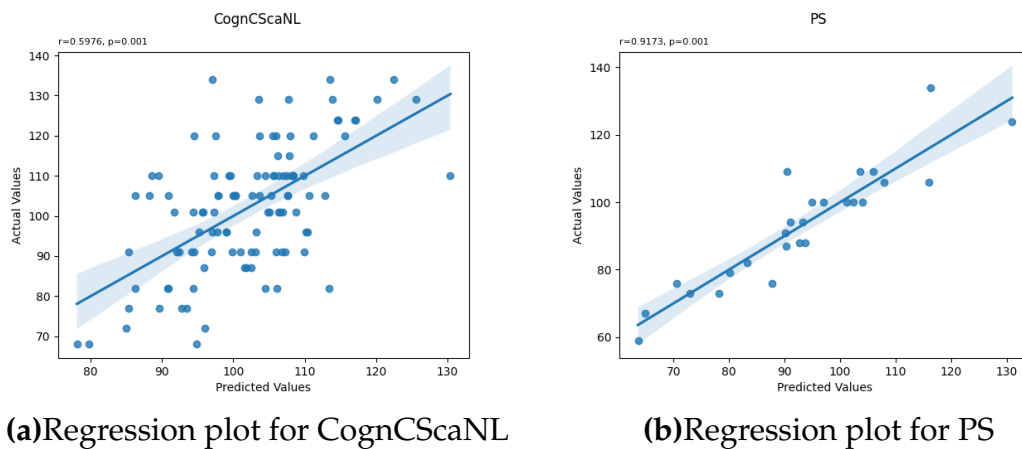
Figure 4.5: Regression results for the combination of aEEG-EEG and MRI for cognitive composite score and processing speed score.

In Figure 4.5 the x-axis shows the predicted values, the y-axis shows the actual values, the solid line indicates linear regression fit, and the shaded area indicates a 95% confidence interval.

It can be seen that all the regression results for the datasets studied are significant. The predictions made mostly show a moderately positive correlation with the actual values. Furthermore, the Fine, Gross motor scaled scores, Total motor composite scores, and all the WPPSI-III scores show a strong positive correlation for all datasets. It can be seen that the regression model that combines both aEEG-EEG and MRI prevails over the individual models.

4.3   Classification Results

A 3-fold weighted logistic regression model was trained on each variation of the datasets. Each fold was stratified to maintain the ratio of instances of each class. Models with at least 2 samples per class per fold were considered robust. Further, performance metrics were estimated with 95% Confidence Intervals (CIs) based on 1000 bootstrap resamples.

Log likelihoods of probability estimates were used as the measure to calculate the AIC and BIC for each model.

		aEEG-EEG		MRI NS	
Target Variable	**Feature Reduction**	**AIC**	**BIC**	**AIC**	**BIC**
CognCScaNL	None	651.172	1843.473	299.859	810.322
	Factor Analysis	81.963	231.568	62.585	167.705
	MRMR	115.145	342.078	75.954	217.566
FMTRS	None	651.352	1831.208	300.085	804.997
	Factor Analysis	82.57	227.622	63.031	166.262
	MRMR	115.785	340.35	76.098	216.171
GMTRS_NLca	None	651.427	1799.355	300.099	787.056
	Factor Analysis	82.364	222.951	63.129	161.946
	MRMR	115.61	334.098	76.125	211.216
TMotCScaNL	None	651.576	1793.87	300.216	784.431
	Factor Analysis	82.655	222.949	63.199	164.144
	MRMR	115.588	333.003	76.369	210.7
CBCL_tot2y	None	651.111	1812.718	299.659	802.13
	Factor Analysis	82.041	224.353	62.464	164.284
	MRMR	115.142	336.234	75.661	215.056
FSIQ	None	652.222	1650.942	301.196	701.866
	Factor Analysis	83.255	196.687	63.771	144.717
	MRMR	116.493	306.582	77.322	188.476
VIQ	None	652.241	1650.961	301.129	701.799
	Factor Analysis	83.242	197.2	63.863	143.765
	MRMR	116.274	306.363	77.338	188.492
PIQ	None	652.441	1651.161	301.029	701.699
	Factor Analysis	83.043	200.035	63.654	144.6
	MRMR	116.191	306.28	77.065	188.219
PS	None	652.363	1651.082	301.142	701.812
	Factor Analysis	83.127	202.035	64.123	140.115
	MRMR	116.447	306.536	77.605	188.759

Table 4.5: AIC/BIC scores for logistic regression (Best Models)

The models with the lowest values for both turned out to be the blanketing Factor Analysis. With aEEG-EEG scoring an average value of 82.696 AIC and 213.933 BIC, and MRI scoring 63.313 AIC and 155.282 BIC.

The performance metrics used for the best models include the balanced accuracy(BA) and the f-beta score.

$$A_{balanced} = \frac{sensitvity + specificity}{2} \tag{4.2}$$

$$f_\beta = (1 + \beta^2) \cdot \frac{precision \cdot recall}{(\beta^2 \cdot precision) + recall} \tag{4.3}$$

$$specificity = \frac{TN}{TN+FP}, \quad sensitivity = \frac{TP}{TP+FN},$$
$$precision = \frac{TP}{TP+FP}, \quad recall = \frac{TP}{TP+FN} \tag{4.4}$$

Where, TP is True Positive, TN is True Negative, FP is False Positive, and FN is False Negative All of these range from 0 (completely incorrect) to 1(completely correct).

4.3.1   aEEG-EEG

aEEG-EEG	BA (95% CI)	p (BA)	Fβ (95% CI)	p (Fβ)	Precision (95% CI)	Recall (95% CI)
BSITD-III						
Cognitive composite	0.729 (0.661-0.795)	<0.001	0.557 (0.401-0.712)	<0.001	0.356 (0.267-0.483)	0.560 (0.4-0.72)
Fine motor scaled	0.718 (0.649-0.789)	<0.001	0.568 (0.428-0.709)	<0.001	0.333 (0.241-0.469)	0.571 (0.429-0.714)
Gross motor scaled	0.704 (0.621-0.792)	<0.001	0.521 (0.315-0.679)	<0.001	0.324 (0.219-0.5)	0.526 (0.316-0.684)
Total motor composite	0.67 (0.603-0.755)	<0.001	0.459 (0.308-0.647)	<0.001	0.343 (0.245-0.474)	0.462 (0.308-0.654)
CBCL						
Total behavioural problem	0.637 (0.567-0.71)	0.014	0.287 (0.144-0.430)	0.013	0.4 (0.167-1.0)	0.286 (0.143-0.429)
WPPSI-III						
Full-scale IQ	0.670 (0.589-0.773)	0.005	0.375 (0.249-0.619)	0.006	0.375 (0.182-1.0)	0.375 (0.25-0.625)
Verbal IQ	0.666 (0.589-0.766)	0.003	0.375 (0.249-0.617)	0.006	0.375 (0.167-0.75)	0.375 (0.25-0.625)
Performance IQ*	0.689 (0.593-0.793)	0.013	0.429 (0.222-1.0)	0.017	0.429 (0.222-1.0)	0.4 (0.2-0.6)
Processing speed	0.652 (0.584-0.724)	0.008	0.334 (0.25-0.5)	0.006	0.357 (0.217-0.601)	0.333 (0.25-0.5)

Table 4.6: Prediction performance of logistic regression on aEEG-EEG

The prediction performance of the aEEG models was moderately high for BSITD-III (BA=$0.67 - 0.729$), and moderate for CBCL (BA=0.637) and WPPSI-III (BA=$0.652 - 0.689$). For the latter, they also had low Fβ scores (CBCL Fβ=0.287, WPPSI-III Fβ=$0.334 - 0.429$). Performance IQ is starred due to insufficient instances in each class during resampling, consequently rendering its performance metrics unverifiable.

(a) Confusion Matrix for CognCScaNL

(b) Confusion Matrix for PS

(c) SHAP values for CognCScaNL
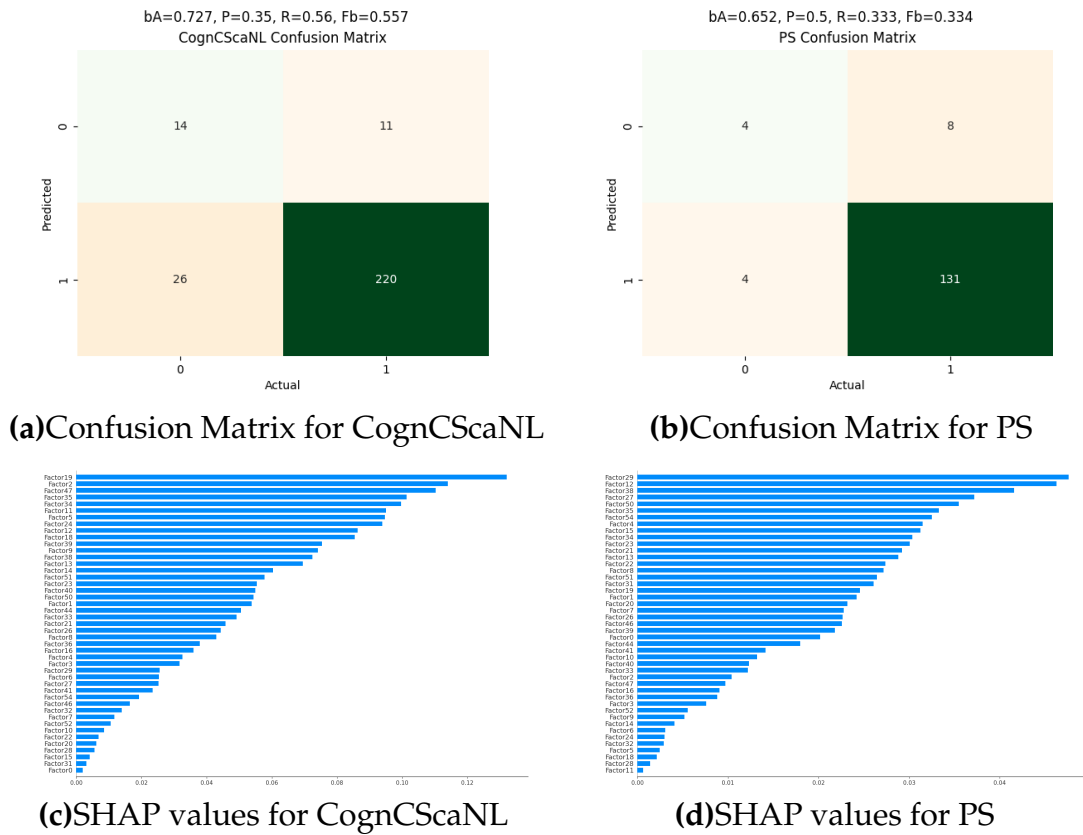
(d) SHAP values for PS

Figure 4.6: Classification results for aEEG-EEG for cognitive composite score and processing speed score

Examining the top five SHAP values for the factors most influential to the results, the following is observed[4]:

- *Cognition composite score:* Factor 19, Factor 2, Factor 47, Factor 35 and Factor 34, correspond to a linear fusion of statistics of aEEG features involving amplitude (day 1), range/amplitude (day 1), amplitude (day 2), range (day 2), and amplitude (day 3), respectively.

- *Processing speed score:* Factor 29, Factor 12, Factor 38, Factor 27 and Factor 50, correspond to a linear fusion of statistics of aEEG features involving spectral content (day 2), spectral content (day 2), connectivity (day 2), spectral content (day 3), and spectral content (day 2), respectively.

---

[4] All features corresponding to the factor are present in Appendix D.

4.3.2 MRI

MRI NS	BA (95% CI)	p (BA)	Fβ (95% CI)	p (Fβ)	Precision (95% CI)	Recall (95% CI)
BSITD-III						
Cognitive composite	0.724 (0.654-0.802)	<0.001	0.523 (0.381-0.666)	<0.001	0.446 (0.318-0.625)	0.524 (0.381-0.666)
Fine motor scaled	0.746 (0.663-0.818)	<0.001	0.642 (0.448-0.79)	<0.001	0.361 (0.256-0.5)	0.65 (0.45-0.8)
Gross motor scaled	0.727 (0.65-0.798)	<0.001	0.532 (0.334-0.664)	<0.001	0.444 (0.286-0.643)	0.533 (0.333-0.667)
Total motor composite	0.707 (0.636-0.779)	<0.001	0.524 (0.368-0.679)	<0.001	0.412 (0.273-0.588)	0.526 (0.368-0.684)
CBCL						
Total behavioural problem*	0.672 (0.58-0.694)	0.006	0.395 (0.199-0.401)	<0.001	0.333 (0.125-1.0)	0.4 (0.2-0.4)
WPPSI-III						
Full-scale IQ*	0.683 (0.589-0.783)	0.025	0.4 (0.201-0.599)	0.012	0.5 (0.25-1.0)	0.4 (0.2-0.6)
Verbal IQ*	0.694 (0.589-0.833)	0.019	0.402 (0.201-0.783)	0.017	0.667 (0.3-1.0)	0.4 (0.2-0.8)
Performance IQ*	0.666 (0.65-0.667)	0.096	0.336 (0.332-0.336)	0.024	1.0 (0.25-1.0)	0.333 (0.333-0.333)
Processing speed	0.805 (0.650-0.894)	0.002	0.663 (0.334-0.829)	0.002	0.5 (0.25-1.0)	0.667 (0.333-0.833)

Table 4.7: Prediction performance of logistic regression on MRI (Unscaled)

The prediction performance of the MRI models was slightly better than that of aEEG-EEG for eligible outcomes in BSITD-III (BA=$0.707 - 0.746$, fβ=$0.523 - 0.642$) and WPPSI-III (BA=0.805, Fβ=0.663). Total behavioural problem, Full-scale IQ, Verbal IQ and Performance IQ are starred due to a lack of instances in each class when resampling occurred. These metrics are not verifiable.

(a) Confusion Matrix for CognCScaNL

(b) Confusion Matrix for PS

(c) SHAP values for CognCScaNL
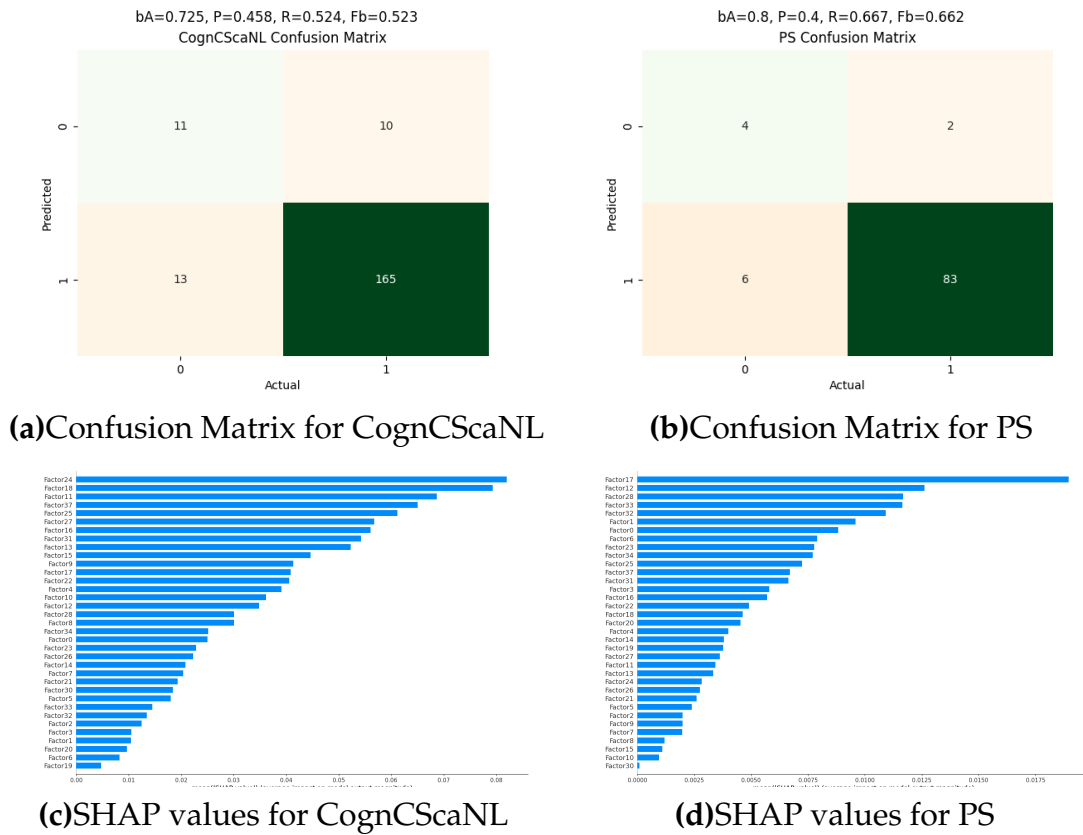
(d) SHAP values for PS

Figure 4.7: Classification results for MRI for cognitive composite score and processing speed score

Examining the top five SHAP values for the factors most influential to the results, the following is observed[5]:

- *Cognition composite score:* Factor 24, Factor 18, Factor 11, Factor 37 and Factor 25, correspond to a linear fusion of volumes of medial and inferior temporal gyri WM, Kidokoro cerebellum overall score, thalamus (low intensity), entire brain, and anterior temporal lobe GM, respectively.

- *Processing speed score:* Factor 17, Factor 12, Factor 28, Factor 33 and Factor 32, correspond to a linear fusion of volumes of gyri parahippocampalis GM, brainstem, cingulate gyrus GM, Cingulate gyrus WM, and Cingulate gyrus WM, respectively.

---

[5] All features corresponding to the factor are present in Appendix D.

4.3.3 Combination

Combination	BA (95% CI)	p (BA)	Fβ (95% CI)	p (Fβ)	Precision (95% CI)	Recall (95% CI)
BSITD-III						
Cognitive composite	0.894	<0.001	0.876	<0.001	0.517	0.882
Fine motor scaled	0.856	<0.001	0.836	<0.001	0.471	0.842
Gross motor scaled	0.832	<0.001	0.779	<0.001	0.423	0.786
Total motor composite	0.898	<0.001	0.878	<0.001	0.6	0.882
CBCL						
Total behavioural problem*	0.872	0.006	0.75	0.004	0.75	0.75
WPPSI-III						
Full-scale IQ	0.907	<0.001	0.855	<0.001	0.667	0.857
Verbal IQ	0.9	<0.001	0.854	<0.001	0.6	0.857
Performance IQ	0.9	0.002	0.801	0.003	1.0	0.8
Processing speed	0.795	<0.001	0.742	<0.001	0.353	0.75

Table 4.8: Prediction performance of logistic regression on the combination of aEEG-EEG and MRI (Unscaled)

The prediction performance of the combination model was better than that of aEEG or MRI alone. BSITD-III: BA=0.832 − 0.898, Fβ=0.779 − 0.878; WPPSI-III: BA=0.795 − 0.907, Fβ=0.742 − 0.855.

(a) Confusion Matrix for CognCScaNL

(b) Confusion Matrix for PS
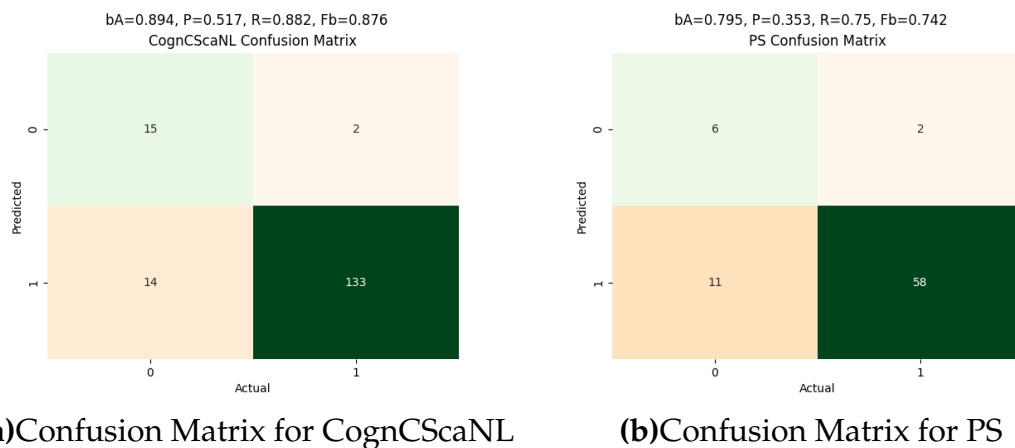
Figure 4.8: Classification results for the combination of aEEG-EEG and MRI for cognitive composite score and processing speed score

It is important to note that the evaluation of this combined model did not adhere to the same protocol employed for the aEEG and MRI models. The split of training and testing data for the former two models and the combined model occurred independently to assess the performance of the combined model. This approach was necessitated by the absence of any intersecting samples in the test sets between aEEG and MRI.

5. Discussion

This thesis delved into the potential of aEEG and MRI in predicting neurodevelopmental outcomes for extremely preterm infants.

Building upon the work done by Wang et al., this research was able to find configurations of aEEG-EEG data that led to better predictive performance [12]. Moreover, the addition of MRI allowed for a comparative analysis with another modality of neurological input which ushered the development of new prediction models with even better prediction results.

Through the employment of multiple feature reduction techniques, it was determined that factor analysis was the most optimal. It uses factor loadings to identify underlying variables that can elucidate variances in the dataset, hence reducing dimensionality. The results yielded were found to have better performance when compared to those without any feature reduction, while substantially reducing model complexity.

In addition, results showed that analyzing MRI alone led to more correct predictions for BSITD-III scores, especially for those associated with motor function. MRI also showed slightly better results for CBCL and WPPSI-III scores for linear regression.

Similarly, the model that combines MRI and aEEG predictions also shows better predictive performance for linear regression. The logistic regression results, however, were calculated on a different split of datasets than those of the aEEG and MRI alone. While they indeed exhibit superior performance, it is important to acknowledge that the results cannot be unequivocally regarded as factual given the present configuration of the data.

Additionally, it was found that disabling any scaling of MRI data availed, comparatively, the best results. This was in contrast to the notion that as volumes in the brain would increase as an infant grows, the scoring of the brain impairment would need a flat volume slope to distinguish between infants

with abnormalities and those without. This is an assumption that the wider body at the UMCU is trying to challenge.

Looking at the SHAP values that explain the models tested, it is evident that the cognitive composite score is derived from values of the range of the aEEG, while that for the processing speed quotient comes from the aEEG features involving spectral content. Due to the novelty of the research conducted, currently literature does not exist for the further verification of the results shown by the SHAP values for aEEG predictions.

For the explanation of features regarding predictions made using the MRI dataset, a comparison is made with the works of Inder et al. and Kidokoro's conclusions [44] [43]. Linear regression of the MRI dataset shows values derived from GM, deep GM, CSF and the temporal lobe for the cognitive composite score, and from GM, WM, and temporal lobe for the processing speed quotient. These findings are mostly consistent with Inder's work. Logistic regression, however, points additionally to the thalamus as a source of influence for cognitive abilities. This has been regarded as a part of the brain that has no clear evidence of association to the cognitive function.

Despite initially appearing to encompass a sizable cohort of infants to study and analyse, the mismatch of the several datasets causes a scarcity of usable information. A major loss stems from the unavailability of an independent evaluation set that could determine the truthful performance of the models with real-world values. This can only be overcome by the collection of more data.

The contributions of features to the models were evaluated through SHAP values. It is important to note that SHAP values solely assess the significance of features within the particular model, rather than its broader real-world importance. Considering the potential for inaccuracies in predictions made by the machine learning models, SHAP values may not consistently reflect reality. Future research should consider methodologies that incorporate an assessment of features' real-world importance with validation from an external source.

Furthermore, the prognosis of an extremely preterm infant does not solely

hinge on the data collected within a limited temporal snapshot. A reliable prediction would ideally include a progressive timeline of clinical and neuroimaging data, enabling the model to differentiate circumstances that deviate from that of a healthy infant.

Ultimately, in the future, a state-of-the-art machine learning model may be able to predict neurodevelopmental abnormalities with higher certainties. This would aid a medical professional by allowing them to administer precision treatments without wasting time scouring through the massive amounts of medical data that can be collected.

6. Conclusion

This study establishes that MRI datasets may not need to be heavily pre-processed by scaling volumes with respect to age, or relative to specific volumes. Moreover, models can be trained on a limited number of features, and an even smaller number of factors to suitably predict long term outcomes of neurological developments. The findings made here offer the prospect of developing an automated tool for long-term neurodisability prognosis, which could facilitate early precision medicine interventions in the Neonatal Intensive Care Unit.

Bibliography

[1] B. J. Stoll, N. I. Hansen, E. F. Bell, *et al.*, "Trends in care practices, morbidity, and mortality of extremely preterm neonates, 1993-2012," *JAMA*, vol. 314, pp. 1039–1051, 10 Sep. 2015, ISSN: 0098-7484. DOI: 10.1001/JAMA.2015.10244. [Online]. Available: https://jamanetwork.com/journals/jama/fullarticle/2434683.

[2] K. L. Costeloe, E. M. Hennessy, S. Haider, F. Stacey, N. Marlow, and E. S. Draper, "Short term outcomes after extreme preterm birth in england: Comparison of two birth cohorts in 1995 and 2006 (the epicure studies)," *BMJ (Clinical research ed.)*, vol. 345, 7886 Dec. 2012, ISSN: 1756-1833. DOI: 10.1136/BMJ.E7976. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/23212881/.

[3] T. M. Berger, M. A. Steurer, A. Woerner, P. Meyer-Schiffer, and M. Adams, "Trends and centre-to-centre variability in survival rates of very preterm infants (<32 weeks) over a 10-year-period in switzerland," *Archives of Disease in Childhood - Fetal and Neonatal Edition*, vol. 97, F323–F328, 5 Sep. 2012, ISSN: 1359-2998. DOI: 10.1136/FETALNEONATAL-2011-301008. [Online]. Available: https://fn.bmj.com/content/97/5/F323%20https://fn.bmj.com/content/97/5/F323.abstract.

[4] H. C. Glass, A. T. Costarino, S. A. Stayer, C. M. Brett, F. Cladis, and P. J. Davis, "Outcomes for extremely premature infants," *Anesthesia and analgesia*, vol. 120, pp. 1337–1351, 6 Jun. 2015, ISSN: 1526-7598. DOI: 10.1213/ANE.0000000000000705. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/25988638/.

[5] H. Nakanishi, H. Suenaga, A. Uchiyama, Y. Kono, and S. Kusuda, "Trends in the neurodevelopmental outcomes among preterm infants from 2003-2012: A retrospective cohort study in japan," *Journal of perinatology : official journal of the California Perinatal Association*, vol. 38, pp. 917–928, 7 Jul. 2018, ISSN: 1476-5543. DOI: 10.1038/S41372-018-0061-7. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/29679045/.

[6] E. S. Twilhaar, R. M. Wade, J. F. D. Kieviet, J. B. V. Goudoever, R. M. V. Elburg, and J. Oosterlaan, "Cognitive outcomes of children born extremely or very preterm since the 1990s and associated risk factors: A meta-analysis and meta-regression," *JAMA pediatrics*, vol. 172, pp. 361–367, 4 Apr. 2018, ISSN: 2168-6211. DOI: 10.1001/JAMAPEDIATRICS.2017.5323. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/29459939/.

[7] E. E. Rogers and S. R. Hintz, "Early neurodevelopmental outcomes of extremely preterm infants," *Seminars in perinatology*, vol. 40, pp. 497–

509, 8 Dec. 2016, ISSN: 1558-075X. DOI: 10.1053/J.SEMPERI.2016.09.002. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/27865437/.

[8] V. Pierrat, L. Marchand-Martin, C. Arnaud, *et al.*, "Neurodevelopmental outcome at 2 years for preterm children born at 22 to 34 weeks' gestation in france in 2011: Epipage-2 cohort study," *BMJ (Clinical research ed.)*, vol. 358, 2017, ISSN: 1756-1833. DOI: 10.1136/BMJ.J3448. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/28814566/.

[9] *Inbrief: The science of early childhood development.* [Online]. Available: https://developingchild.harvard.edu/resources/inbrief-science-of-ecd/.

[10] J. J. Volpe, "Dysmaturation of premature brain: Importance, cellular mechanisms, and potential interventions," *Pediatric neurology*, vol. 95, pp. 42–66, Jun. 2019, ISSN: 1873-5150. DOI: 10.1016/J.PEDIATRNEUROL.2019.02.016. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/30975474/.

[11] M. L. Tataranno, D. C. Vijlbrief, J. Dudink, and M. J. Benders, "Precision medicine in neonates: A tailored approach to neonatal brain injury," *Frontiers in Pediatrics*, vol. 9, p. 418, May 2021, ISSN: 22962360. DOI: 10.3389/FPED.2021.634092/BIBTEX.

[12] X. Wang, C. Trabatti, J. D. Lauren Weeke, *et al.*, "Early postnatal (a)eeg features predict long-term neurodevelopmental outcomes in extremely preterm infants a 10-year cohort study," 2023.

[13] C. van 't Westende, V. J. Geraedts, T. van Ramesdonk, *et al.*, "Neonatal quantitative electroencephalography and long-term outcomes: A systematic review," *Developmental medicine and child neurology*, vol. 64, pp. 413–420, 4 Apr. 2022, ISSN: 1469-8749. DOI: 10.1111/DMCN.15133. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/34932822/.

[14] E. Bourel-Ponchel, S. Gueden, D. Hasaerts, *et al.*, "Normal eeg during the neonatal period: Maturational aspects from premature to full-term newborns," *Neurophysiologie clinique = Clinical neurophysiology*, vol. 51, pp. 61–88, 1 Jan. 2021, ISSN: 1769-7131. DOI: 10.1016/J.NEUCLI.2020.10.004. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/33239230/.

[15] N. M. E. Ters, Z. A. Vesoulis, S. M. Liao, C. D. Smyser, and A. M. Mathur, "Term-equivalent functional brain maturational measures predict neurodevelopmental outcomes in premature infants," *Early human development*, vol. 119, p. 68, Apr. 2018, ISSN: 18726232. DOI: 10.1016/J.EARLHUMDEV.2018.02.010. [Online]. Available: /pmc/articles/PMC6190680/%20/pmc/articles/PMC6190680/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6190680/.

[16] E. P. Fogtmann, A. M. Plomgaard, G. Greisen, and C. Gluud, "Prognostic accuracy of electroencephalograms in preterm infants: A sys-

tematic review," *Pediatrics*, vol. 139, 2 Feb. 2017, ISSN: 1098-4275. DOI: `10.1542/PEDS.2016-1951`. [Online]. Available: `https://pubmed.ncbi.nlm.nih.gov/28143915/`.

[17] J. Dubois, M. Alison, S. J. Counsell, L. Hertz-Pannier, P. S. Hüppi, and M. J. Benders, "Mri of the neonatal brain: A review of methodological challenges and neuroscientific advances," *Journal of Magnetic Resonance Imaging*, vol. 53, p. 1318, 5 May 2021, ISSN: 15222586. DOI: `10.1002/JMRI.27192`. [Online]. Available: `/pmc/articles/PMC8247362/%20/pmc/articles/PMC8247362/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8247362/`.

[18] *Brain anatomy | mayfield brain & spine cincinnati, ohio*. [Online]. Available: `https://mayfieldclinic.com/pe-anatbrain.htm`.

[19] N. A. Shah and C. J. Wusthoff, "How to use: Amplitude-integrated eeg (aeeg)," *Archives of disease in childhood. Education and practice edition*, vol. 100, pp. 75–81, 2 Apr. 2015, ISSN: 1743-0593. DOI: `10.1136/ARCHDISCHILD-2013-305676`. [Online]. Available: `https://pubmed.ncbi.nlm.nih.gov/25035312/`.

[20] M. C. Toet and P. M. Lemmers, "Brain monitoring in neonates," *Early Human Development*, vol. 85, pp. 77–84, 2 Feb. 2009, ISSN: 0378-3782. DOI: `10.1016/J.EARLHUMDEV.2008.11.007`.

[21] F. Doandes, A.-M. Manea, N. Lungu, *et al.*, "Clinical, biological and electroencephalographic monitoring of newborns with neurological risk in the neonatal intensive care unit," *Experimental and therapeutic medicine*, vol. 22, 1 May 2021, ISSN: 1792-0981. DOI: `10.3892/ETM.2021.10192`. [Online]. Available: `https://pubmed.ncbi.nlm.nih.gov/34035857/`.

[22] A. L. Numis and R. A. Shellhaas, "Neonatal seizure management: What is timely treatment and does it influence neurodevelopment?" *The Journal of Pediatrics*, vol. 243, pp. 7–8, Apr. 2022, ISSN: 0022-3476. DOI: `10.1016/J.JPEDS.2021.12.004`.

[23] C. Burger, M. Biermayr, A. Posod, *et al.*, "Amplitude-integrated electroencephalography shows alterations in children born preterm displaying poor literacy precursor skills," *Acta Paediatrica (Oslo, Norway : 1992)*, vol. 108, p. 1661, 9 2019, ISSN: 16512227. DOI: `10.1111/APA.14755`. [Online]. Available: `/pmc/articles/PMC6767598/%20/pmc/articles/PMC6767598/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6767598/`.

[24] C. Burger, M. Hammerl, V. Neubauer, U. P. Peglow, U. Kiechl-Kohlendorfer, and E. Griesmaier, "Early preterm infants with abnormal psychomotor neurodevelopmental outcome at age two show alterations in amplitude-integrated electroencephalography signals," *Early human development*, vol. 141, Feb. 2020, ISSN: 1872-6232. DOI: `10.1016/J.EARLHUMDEV.2019.104935`. [Online]. Available: `https://pubmed.ncbi.nlm.nih.gov/31835163/`.

[25] K. K. Iyer, J. A. Roberts, L. Hellström-Westas, *et al.*, "Cortical burst dynamics predict clinical outcome early in extremely preterm in-

fants," *Brain : a journal of neurology*, vol. 138, pp. 2206–2218, Pt 8 Aug. 2015, ISSN: 1460-2156. DOI: 10.1093/BRAIN/AWV129. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/26001723/.

[26] A. H. Kong, M. M. Lai, S. Finnigan, R. S. Ware, R. N. Boyd, and P. B. Colditz, "Background eeg features and prediction of cognitive outcomes in very preterm infants: A systematic review," *Early human development*, vol. 127, pp. 74–84, Dec. 2018, ISSN: 1872-6232. DOI: 10.1016/J.EARLHUMDEV.2018.09.015. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/30340071/.

[27] K. Klebermass, M. Olischar, T. Waldhoer, R. Fuiko, A. Pollak, and M. Weninger, "Amplitude-integrated eeg pattern predicts further outcome in preterm infants," *Pediatric research*, vol. 70, pp. 102–108, 1 Jul. 2011, ISSN: 1530-0447. DOI: 10.1203/PDR.0B013E31821BA200. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/21436758/.

[28] R. O. Lloyd, J. M. O'Toole, V. Livingstone, P. M. Filan, and G. B. Boylan, "Can eeg accurately predict 2-year neurodevelopmental outcome for preterm infants?" *Archives of Disease in Childhood. Fetal and Neonatal Edition*, vol. 106, p. 535, 5 Sep. 2021, ISSN: 14682052. DOI: 10.1136/ARCHDISCHILD-2020-319825. [Online]. Available: /pmc/articles/PMC8394766/%20/pmc/articles/PMC8394766/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8394766/.

[29] L. A. D. Wispelaere, S. Ouwehand, M. Olsthoorn, *et al.*, "Electroencephalography and brain magnetic resonance imaging in asphyxia comparing cooled and non-cooled infants," *European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society*, vol. 23, pp. 181–190, 1 Jan. 2019, ISSN: 1532-2130. DOI: 10.1016/J.EJPN.2018.09.001. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/30279083/.

[30] D. Paz-Levy, L. Schreiber, O. Erez, *et al.*, "Inflammatory and vascular placental lesions are associated with neonatal amplitude integrated eeg recording in early premature neonates," *PLOS ONE*, vol. 12, e0179481, 6 Jun. 2017, ISSN: 1932-6203. DOI: 10.1371/JOURNAL.PONE.0179481. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0179481.

[31] H. H. Jasper, "Ten-twenty electrode system of the international federation," *Electroencephalography Clinical Neurophysiology*, vol. 10, pp. 371–375, 1958.

[32] A. Plaisier, P. Govaert, M. H. Lequin, and J. Dudink, "Optimal timing of cerebral mri in preterm infants to predict long-term neurodevelopmental outcome: A systematic review," *AJNR: American Journal of Neuroradiology*, vol. 35, p. 841, 5 2014, ISSN: 1936959X. DOI: 10.3174/AJNR.A3513. [Online]. Available: /pmc/articles/PMC7964554/%20/pmc/articles/PMC7964554/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7964554/.

[33] M. A. Rutherford, V. Supramaniam, A. Ederies, *et al.*, "Magnetic resonance imaging of white matter diseases of prematurity," *Neuroradiology*, vol. 52, pp. 505–521, 6 2010, ISSN: 1432-1920. DOI: 10.1007/S00234-010-0700-Y. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/20422407/.

[34] C. D. Smyser, H. Kidokoro, and T. E. Inder, "Mri of the brain at term equivalent age in extremely premature neonates – to scan or not to scan?" *Journal of paediatrics and child health*, vol. 48, p. 794, 9 Sep. 2012, ISSN: 10344810. DOI: 10.1111/J.1440-1754.2012.02535.X. [Online]. Available: /pmc/articles/PMC3595093/%20/pmc/articles/PMC3595093/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3595093/.

[35] A. Plaisier, M. M. Raets, G. M. Ecury-Goossen, *et al.*, "Serial cranial ultrasonography or early mri for detecting preterm brain injury?" *Archives of disease in childhood. Fetal and neonatal edition*, vol. 100, F293–F300, 4 Jul. 2015, ISSN: 1468-2052. DOI: 10.1136/ARCHDISCHILD-2014-306129. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/25637006/.

[36] H. E. Whyte and S. Blaser, "Limitations of routine neuroimaging in predicting outcomes of preterm infants," *Neuroradiology*, vol. 55 Suppl 2, pp. 3–11, SUPPL. 2 2013, ISSN: 1432-1920. DOI: 10.1007/S00234-013-1238-6. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/23955300/.

[37] L. R. Ment, D. Hirtz, and P. S. Hüppi, "Imaging biomarkers of outcome in the developing preterm brain," *The Lancet. Neurology*, vol. 8, pp. 1042–1055, 11 Nov. 2009, ISSN: 1474-4465. DOI: 10.1016/S1474-4422(09)70257-1. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/19800293/.

[38] L. J. Woodward, P. J. Anderson, N. C. Austin, K. Howard, and T. E. Inder, "Neonatal mri to predict neurodevelopmental outcomes in preterm infants," *The New England journal of medicine*, vol. 355, pp. 685–694, 7 Aug. 2006, ISSN: 1533-4406. DOI: 10.1056/NEJMOA053792. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/16914704/.

[39] A. J. Spittle, N. C. Brown, L. W. Doyle, *et al.*, "Quality of general movements is related to white matter pathology in very preterm infants," *Pediatrics*, vol. 121, 5 May 2008, ISSN: 1098-4275. DOI: 10.1542/PEDS.2007-1924. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/18390959/.

[40] L. E. Dyet, N. Kennea, S. J. Counsell, *et al.*, "Natural history of brain lesions in extremely preterm infants studied with serial magnetic resonance imaging from birth and neurodevelopmental assessment," *Pediatrics*, vol. 118, pp. 536–548, 2 Aug. 2006, ISSN: 1098-4275. DOI: 10.1542/PEDS.2005-1866. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/16882805/.

[41] C. R. Pierson, R. D. Folkerth, S. S. Billiards, *et al.*, "Gray matter injury associated with periventricular leukomalacia in the premature

infant," *Acta neuropathologica*, vol. 114, pp. 619–631, 6 Dec. 2007, ISSN: 0001-6322. DOI: 10.1007/S00401-007-0295-5. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/17912538/.

[42] A. Messerschmidt, P. C. Brugger, E. Boltshauser, *et al.*, "Disruption of cerebellar development: Potential complication of extreme prematurity," *AJNR: American Journal of Neuroradiology*, vol. 26, p. 1659, 7 2005, ISSN: 01956108. [Online]. Available: /pmc/articles/PMC7975 176/%20/pmc/articles/PMC7975176/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7975176/.

[43] H. Kidokoro, J. J. Neil, and T. E. Inder, "New mr imaging assessment tool to define brain abnormalities in very preterm infants at term," *AJNR: American Journal of Neuroradiology*, vol. 34, p. 2208, 11 Nov. 2013, ISSN: 01956108. DOI: 10.3174/AJNR.A3521. [Online]. Available: /pmc/articles/PMC4163698/%20/pmc/articles/PMC4163698/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4163698/.

[44] T. E. Inder, J. J. Volpe, and P. J. Anderson, "Defining the neurologic consequences of preterm birth," *New England Journal of Medicine*, vol. 389, pp. 441–453, 5 Aug. 2023, ISSN: 0028-4793. DOI: 10.1056/NEJMRA2303347/SUPPL_FILE/NEJMRA2303347_DISCLOSURES.PDF. [Online]. Available: https://www.nejm.org/doi/full/10.1056/NEJMra2303347.

[45] M. J. Brouwer, K. J. Kersbergen, B. J. V. Kooij, *et al.*, "Preterm brain injury on term-equivalent age mri in relation to perinatal factors and neurodevelopmental outcome at two years," *PLoS ONE*, vol. 12, 5 May 2017, ISSN: 19326203. DOI: 10.1371/JOURNAL.PONE.0177128. [Online]. Available: /pmc/articles/PMC5423624/%20/pmc/articles/PMC5423624/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5423624/.

[46] C. Wang, K. Allegaert, M. Y. Peeters, D. Tibboel, M. Danhof, and C. A. Knibbe, "The allometric exponent for scaling clearance varies with age: A study on seven propofol datasets ranging from preterm neonates to adults," *British Journal of Clinical Pharmacology*, vol. 77, p. 149, 1 Jan. 2014, ISSN: 03065251. DOI: 10.1111/BCP.12180. [Online]. Available: /pmc/articles/PMC3895356/%20/pmc/articles/PMC3895356/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3895356/.

[47] N. Shrestha, "Factor analysis as a tool for survey analysis," *American Journal of Applied Mathematics and Statistics, Vol. 9, 2021, Pages 4-11*, vol. 9, pp. 4–11, 1 Feb. 2021, ISSN: 2328-7306. DOI: 10.12691/AJAMS-9-1-2. [Online]. Available: http://pubs.sciepub.com/ajams/9/1/2/index.html%20http://pubs.sciepub.com/ajams/9/1/2/abstract.html.

[48] *Properties of sufficiency and statistical tests on jstor*. [Online]. Available: https://www.jstor.org/stable/96803.

[49] D. Bandalos, *Measurement Theory and Applications for the Social Sciences* (Methodology in the Social Sciences). Guilford Publications, 2018, ISBN: 9781462532131. [Online]. Available: `https://books.google.nl/books?id=caxCDwAAQBAJ`.

[50] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," *Journal of bioinformatics and computational biology*, vol. 3, pp. 185–205, 2 Apr. 2005, ISSN: 0219-7200. DOI: `10.1142/S0219720005001004`. [Online]. Available: `https://pubmed.ncbi.nlm.nih.gov/15852500/`.

[51] Z. Zhao, R. Anand, and M. Wang, "Maximum relevance and minimum redundancy feature selection methods for a marketing machine learning platform," *Proceedings - 2019 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2019*, pp. 442–452, Aug. 2019. DOI: `10.1109/DSAA.2019.00059`. [Online]. Available: `https://arxiv.org/abs/1908.05376v1`.

Artificial Intelligence Master Thesis

Brain Imaging Analysis for Prediction of Developmental Problems in Pre-term Infants

(Supplementary Appendix)

First examiner:

Dr. Albert Ali Salah

Second examiner:

Dr. Ronald Poppe

External supervisors:

Dr. Maria Luisa Tataranno

Bob Waalrad

Candidate:

Shreyash Maini

In cooperation with:

UMC Utrecht

Contents

List of Figures

List of Tables

A. Appendix - Distributions

Figure A.1: Distributions of quantitative EEG features across the first three postnatal days

Figure A.1: Day 1, day 2, and day 3 are represented in blue, orange, and green, respectively. 1 = delta frequency band. 2 = theta frequency band. 3 = alpha frequency band. 4 = beta frequency band.
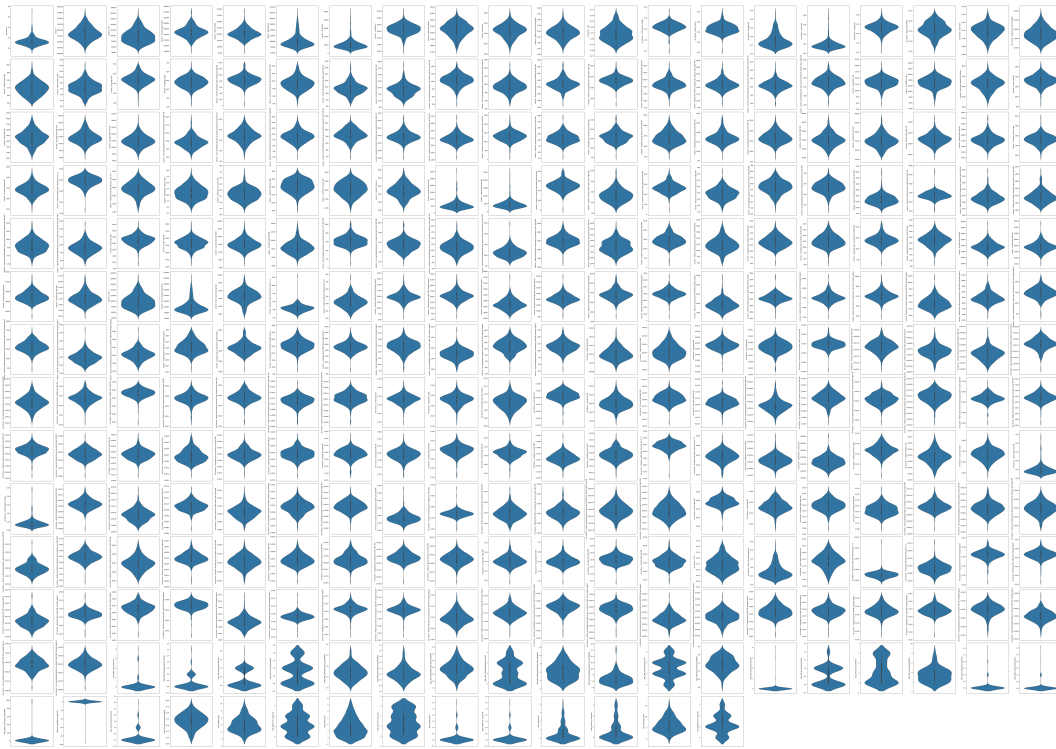
Figure A.2: Distributions of MRI features

B. Appendix - Regression

(a) Cognition   **(b)** Fine Motor   **(c)** Gross Motor

(d) Total Motor   **(e)** Behaviour   **(f)** Full-Scale IQ

(g) Verbal IQ   **(h)** Performance IQ   **(i)** Processing Speed
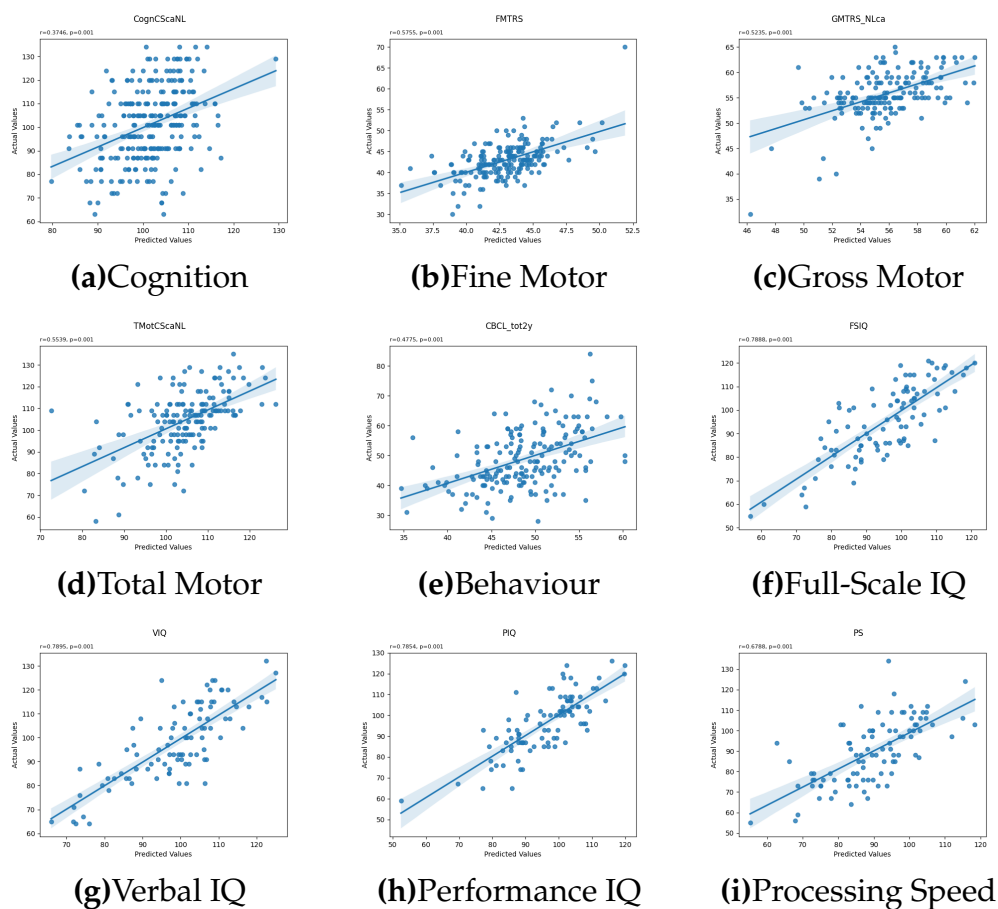
Figure B.1: Linear Regression Results for aEEG-EEG

(a) Cognition

(b) Fine Motor

(c) Gross Motor

(d) Total Motor

(e) Behaviour

(f) Full-Scale IQ

(g) Verbal IQ

(h) Performance IQ

(i) Processing Speed
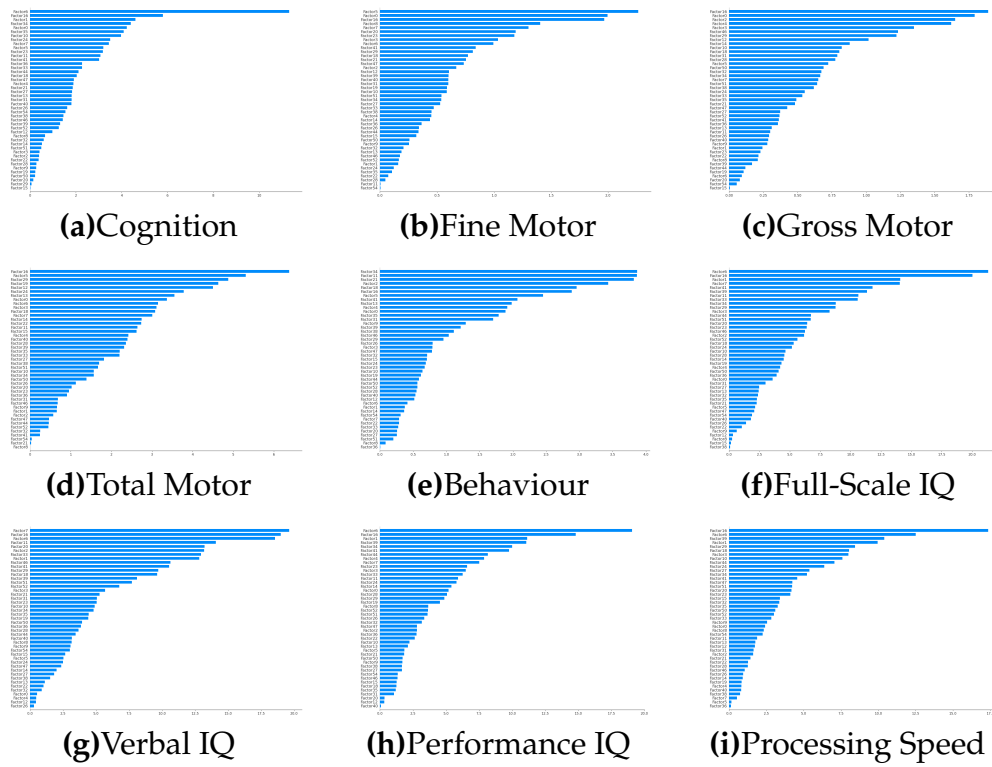
Figure B.2: Linear Regression SHAP graph for aEEG-EEG
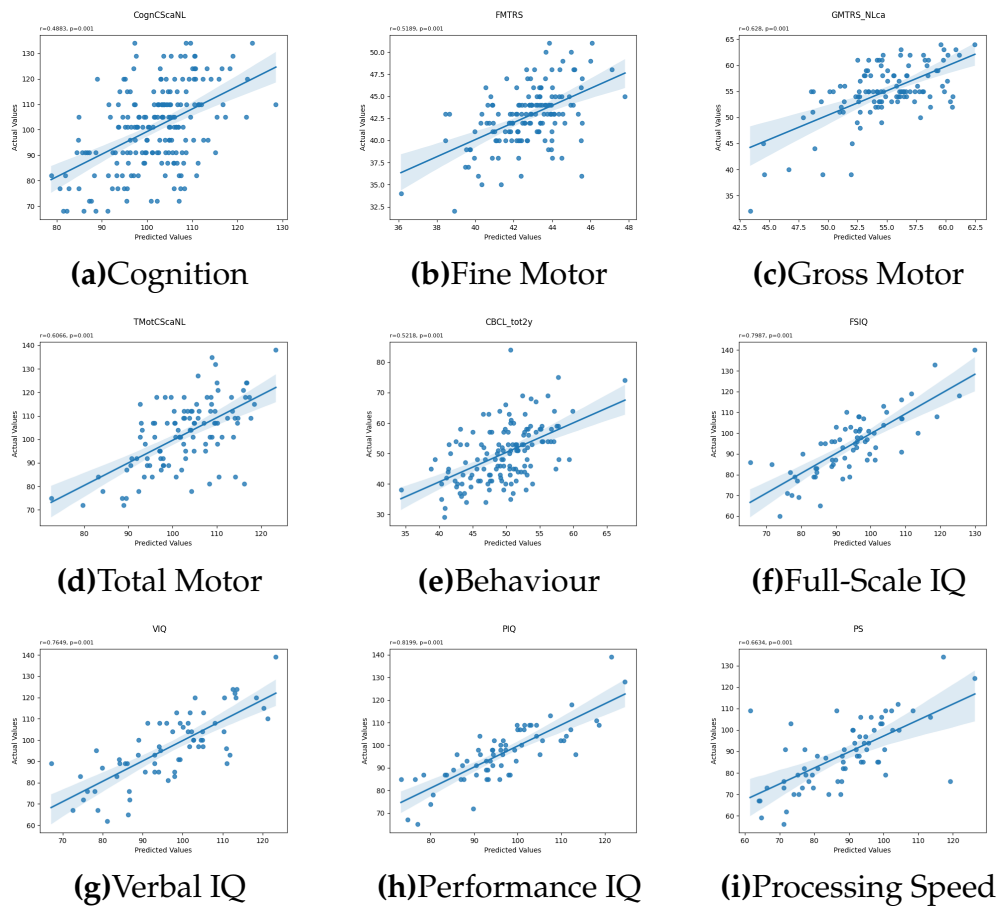
Figure B.3: Linear Regression Results for MRI

(a) Cognition

(b) Fine Motor

(c) Gross Motor

(d) Total Motor

(e) Behaviour

(f) Full-Scale IQ

(g) Verbal IQ

(h) Performance IQ

(i) Processing Speed
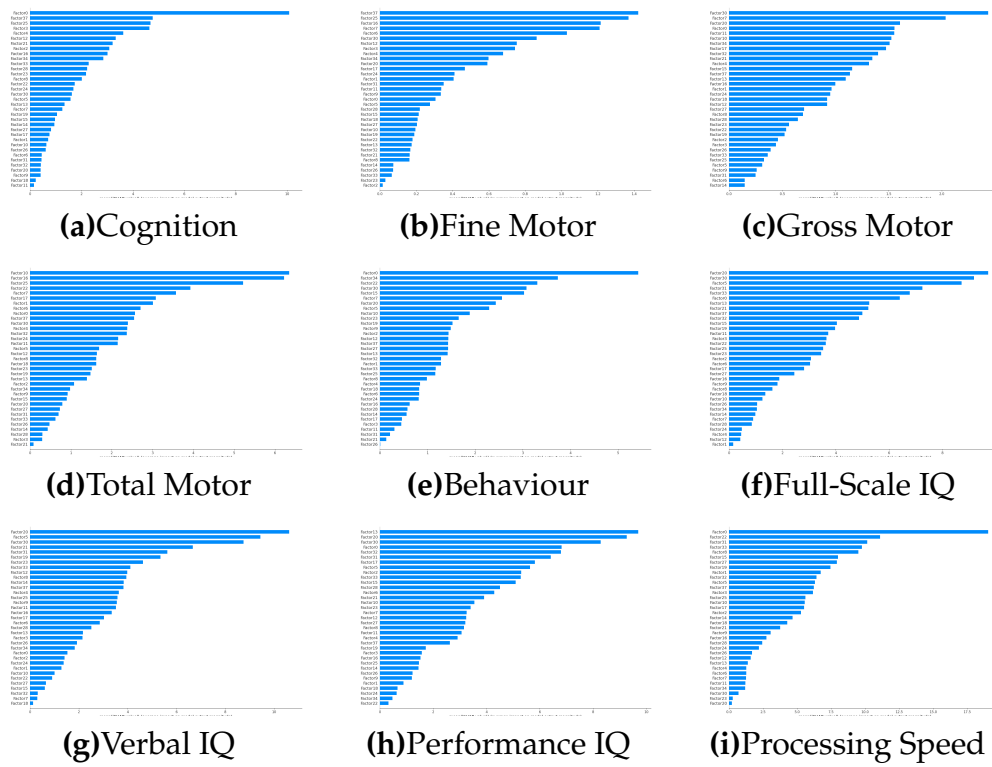
Figure B.4: Linear Regression SHAP graph for MRI

(a) Cognition   **(b)**Fine Motor   **(c)**Gross Motor

(d) Total Motor   **(e)**Behaviour   **(f)**Full-Scale IQ

(g) Verbal IQ   **(h)**Performance IQ   **(i)**Processing Speed
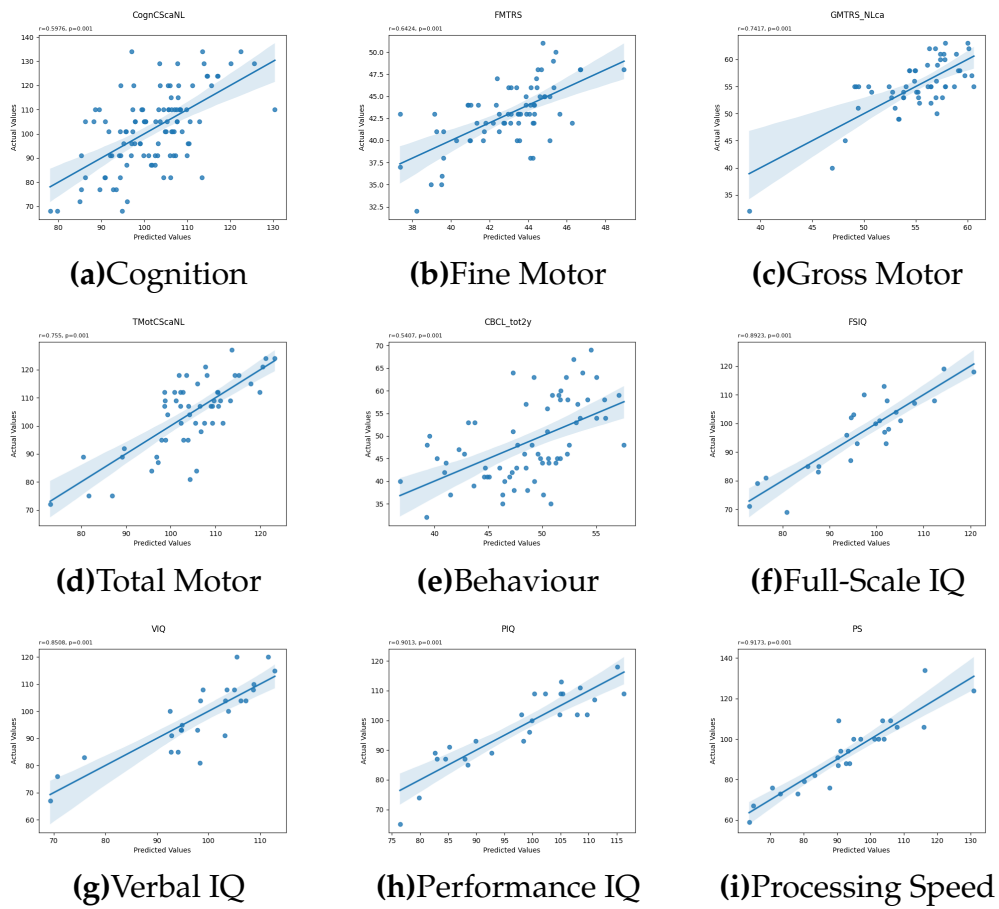
Figure B.5: Linear Regression Results for Combination of aEEG & MRI

C. Appendix - Classification

bA=0.727, P=0.35, R=0.56, Fb=0.557
CognCScaNL Confusion Matrix

(a) Cognition

bA=0.72, P=0.348, R=0.571, Fb=0.568
FMTRS Confusion Matrix

(b) Fine Motor

bA=0.705, P=0.286, R=0.526, Fb=0.522
GMTRS_NLca Confusion Matrix

(c) Gross Motor

bA=0.67, P=0.324, R=0.462, Fb=0.46
TMotCScaNL Confusion Matrix

(d) Total Motor

bA=0.637, P=0.4, R=0.286, Fb=0.287
CBCL_tot2y Confusion Matrix

(e) Behaviour

bA=0.666, P=0.333, R=0.375, Fb=0.375
FSIQ Confusion Matrix

(f) Full-Scale IQ

bA=0.666, P=0.333, R=0.375, Fb=0.375
VIQ Confusion Matrix

(g) Verbal IQ

bA=0.693, P=0.5, R=0.4, Fb=0.401
PIQ Confusion Matrix

(h) Performance IQ

bA=0.652, P=0.5, R=0.333, Fb=0.334
PS Confusion Matrix

(i) Processing Speed

Figure C.1: Classification Results for aEEG-EEG

(a) Cognition **(b)**Fine Motor **(c)**Gross Motor

(d) Total Motor **(e)**Processing Speed
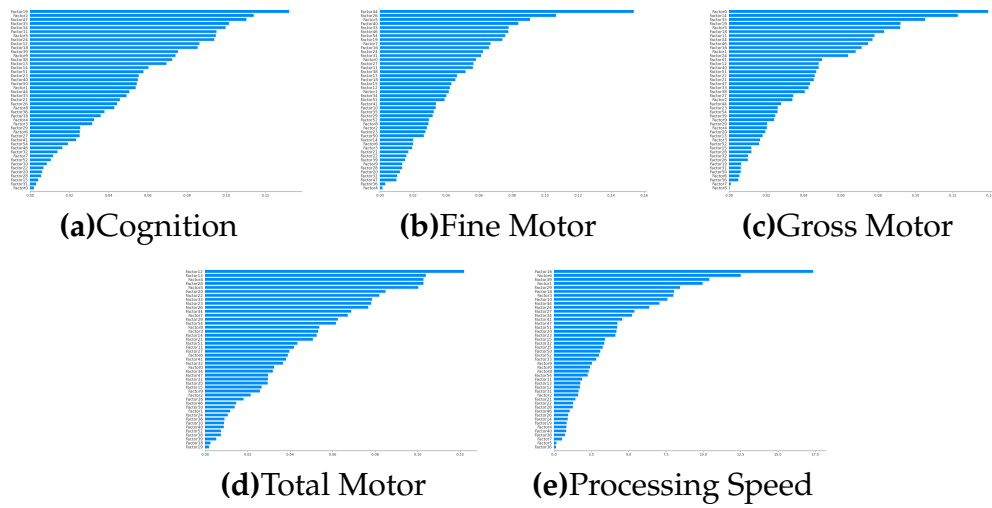
Figure C.2: Classification SHAP graph for aEEG-EEG

FSIQ, VIQ, PIQ, and CBCL scores did not have enough instances in each class when testing the models. Hence, their SHAP graphs have not been included.

(a) Cognition **(b)**Fine Motor **(c)**Gross Motor

(d) Total Motor **(e)**Behaviour **(f)**Full-Scale IQ

(g) Verbal IQ **(h)**Performance IQ **(i)**Processing Speed
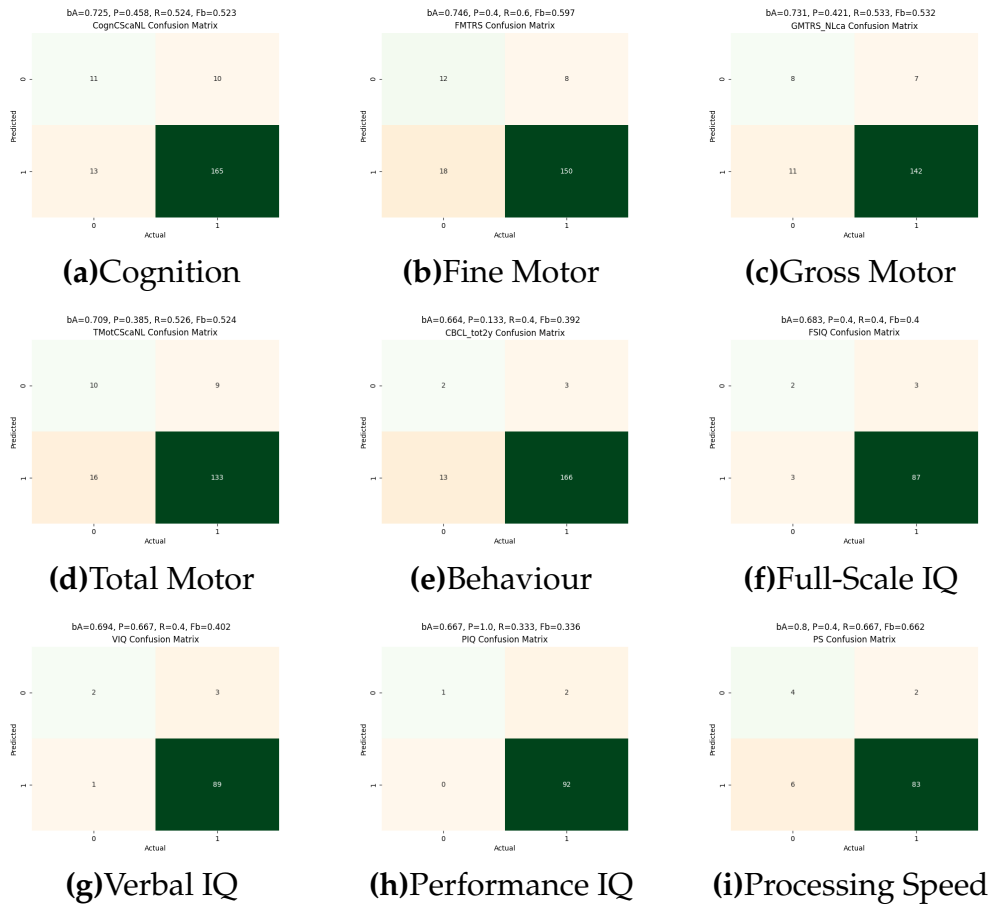
Figure C.3: Classification Results for MRI NS

(a) Cognition **(b)**Fine Motor **(c)**Gross Motor

(d) Total Motor **(e)**Processing Speed
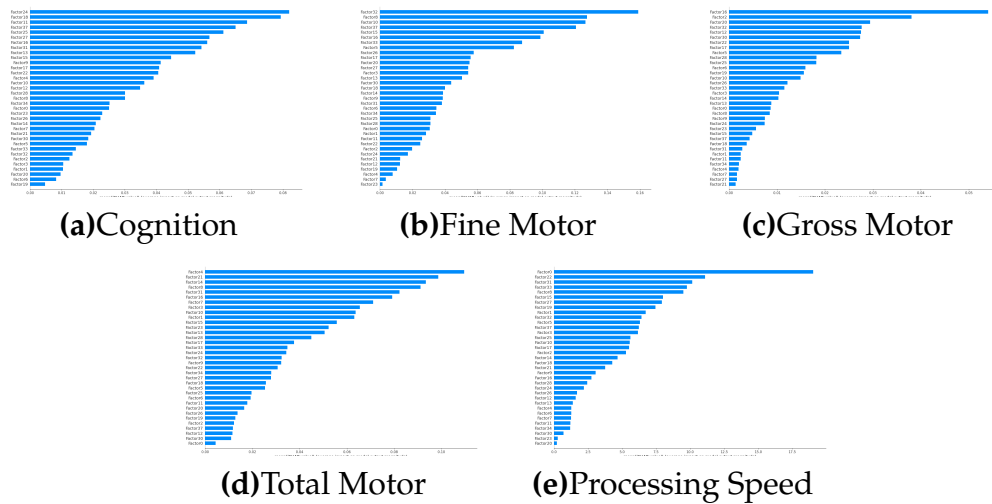
Figure C.4: Classification SHAP graph for MRI NS

FSIQ, VIQ, PIQ, and CBCL scores did not have enough instances in each class when testing the models. Hence, their SHAP graphs have not been

included.

(a) Cognition   **(b)**Fine Motor   **(c)**Gross Motor

(d) Total Motor   **(e)**Behaviour   **(f)**Full-Scale IQ

(g) Verbal IQ   **(h)**Performance IQ   **(i)**Processing Speed
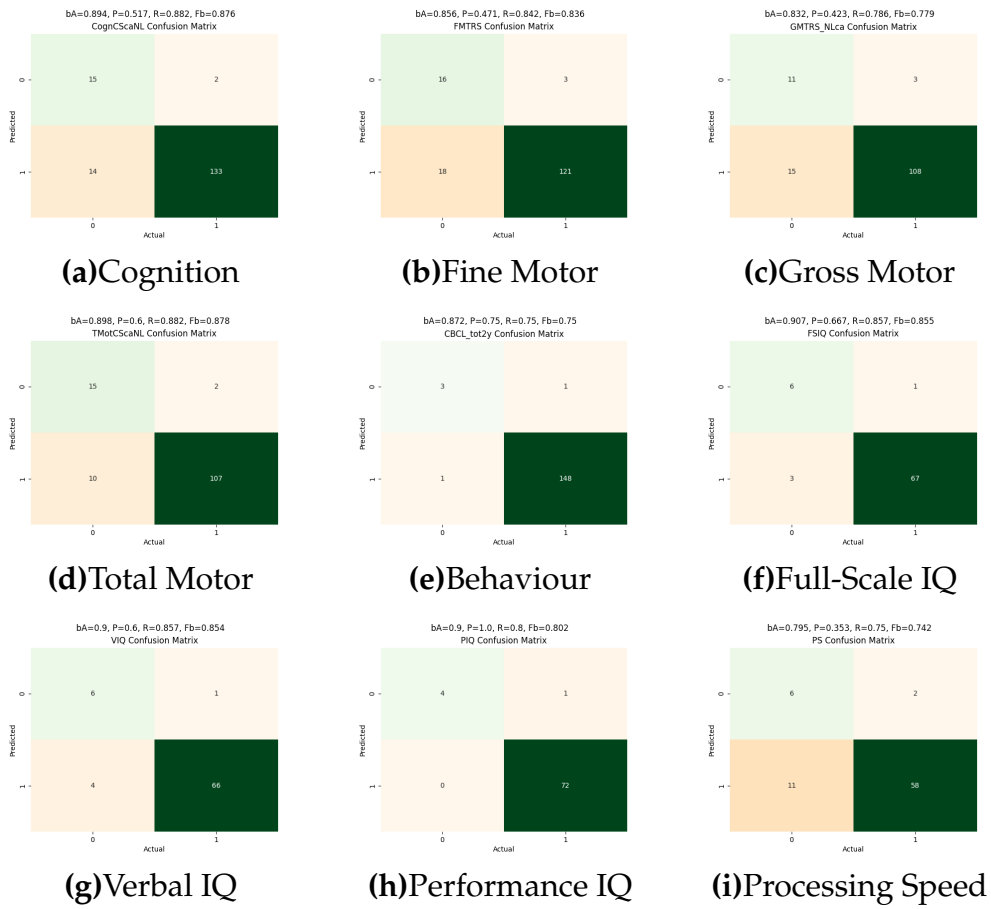
Figure C.5: Classification Results for Combination of aEEG & MRI

D. Appendix - Factor Loadings

Factor	Features
Factor 0	['second_connectivity_coh_max_1', 'second_connectivity_coh_-max_2', 'third_connectivity_corr_3', 'first_asi_hemi', 'third_-connectivity_corr_4', 'third_connectivity_coh_max_2', 'third_-connectivity_coh_mean_1', 'second_connectivity_coh_max_4', 'third_connectivity_coh_max_3', 'third_connectivity_corr_2', 'second_asi_hemi', 'third_asi_hemi', 'second_connectivity_-coh_mean_1', 'third_connectivity_coh_mean_3', 'second_con-nectivity_coh_max_3', 'third_connectivity_coh_mean_4', 'sec-ond_connectivity_BSI_2', 'third_connectivity_coh_mean_2', 'third_connectivity_corr_1', 'third_connectivity_coh_max_1', 'second_connectivity_BSI_3', 'third_connectivity_BSI_3', 'sec-ond_connectivity_coh_mean_2', 'second_connectivity_corr_1', 'second_connectivity_coh_mean_3', 'first_connectivity_BSI_4', 'second_connectivity_coh_mean_4', 'first_connectivity_coh_-max_2', 'second_connectivity_corr_3']
Factor 1	['second_rEEG_asymmetry_4', 'second_rEEG_lower_margin_-1', 'second_rEEG_median_3', 'second_rEEG_CV_2', 'second_-rEEG_asymmetry_3', 'second_rEEG_asymmetry_1', 'second_-rEEG_CV_1', 'second_rEEG_lower_margin_2', 'second_ISP_-Median', 'second_rEEG_median_4', 'second_rEEG_CV_3', 'sec-ond_amplitude_kurtosis_2', 'second_rEEG_median_2']
Factor 2	['first_amplitude_env_SD_1', 'first_amplitude_env_mean_1', 'first_amplitude_total_power_1', 'first_rEEG_upper_margin_1', 'first_rEEG_mean_1', 'first_rEEG_width_1', 'first_rEEG_SD_-1', 'first_amplitude_SD_1', 'first_sp_abs_D', 'first_IBI_burst_-number']

Factor	Features
Factor 3	['first_rEEG_mean_2', 'first_amplitude_SD_2', 'first_rEEG_-mean_3', 'first_sp_abs_A', 'first_rEEG_SD_3', 'first_amplitude_-env_mean_3', 'first_amplitude_total_power_3', 'first_rEEG_-upper_margin_3', 'first_rEEG_width_3', 'first_amplitude_env_-SD_3']
Factor 4	['third_rEEG_asymmetry_2', 'third_rEEG_median_2', 'third_-rEEG_CV_2', 'third_rEEG_asymmetry_4', 'third_rEEG_CV_-1', 'third_rEEG_asymmetry_1', 'third_rEEG_asymmetry_3', 'third_amplitude_kurtosis_3', 'third_rEEG_CV_3']
Factor 5	['third_amplitude_env_SD_2', 'third_rEEG_upper_margin_-2', 'third_sp_abs_T', 'third_rEEG_SD_2', 'third_amplitude_-env_mean_2', 'third_amplitude_total_power_2', 'third_rEEG_-width_2', 'third_amplitude_SD_2']
Factor 6	['first_mse_max', 'first_rEEG_CV_4', 'first_IBI_burst_prc', 'first_rEEG_asymmetry_1', 'first_rEEG_CV_3', 'first_rEEG_-CV_1']
Factor 7	['first_amplitude_env_mean_4', 'first_amplitude_total_power_-4', 'first_rEEG_mean_4', 'first_rEEG_median_4', 'first_ampli-tude_SD_4']
Factor 8	['second_rEEG_SD_2', 'second_amplitude_env_mean_2', 'sec-ond_amplitude_total_power_2']
Factor 9	['first_rEEG_SD_2', 'first_rEEG_upper_margin_2', 'first_rEEG_-width_2']
Factor 10	['third_mse_slope_high', 'third_mse_max', 'third_mse_auc']
Factor 11	['second_sp_abs_D', 'second_rEEG_median_1', 'second_rEEG_-mean_1']
Factor 12	['second_spectral_flatness_3', 'second_spectral_entropy_3']
Factor 13	['third_rEEG_width_3', 'third_rEEG_upper_margin_3']
Factor 14	['first_rEEG_lower_margin_3']
Factor 15	['third_sp_rel_B']
Factor 16	['third_rEEG_lower_margin_1', 'third_rEEG_median_1']

Factor	Features
Factor 17	[]
Factor 18	['first_rEEG_median_2', 'first_rEEG_lower_margin_2', 'first_-rEEG_median_3']
Factor 19	['first_amplitude_total_power_2', 'first_amplitude_env_mean_-2']
Factor 20	['first_rEEG_median_1', 'first_ISP_Median']
Factor 21	['first_rEEG_lower_margin_1']
Factor 22	['first_SATRate_Median']
Factor 23	['first_amplitude_skew_1']
Factor 24	['third_rEEG_SD_1', 'third_rEEG_width_1']
Factor 25	[]
Factor 26	['third_amplitude_kurtosis_2']
Factor 27	['third_spectral_diff_1']
Factor 28	['second_amplitude_skew_1']
Factor 29	['second_mse_slope_high']
Factor 30	[]
Factor 31	['third_amplitude_skew_1']
Factor 32	['first_spectral_diff_4']
Factor 33	['first_rEEG_SD_4', 'first_rEEG_width_4']
Factor 34	['third_amplitude_env_SD_3']
Factor 35	['second_rEEG_asymmetry_2']
Factor 36	['first_IBI_length_median']
Factor 37	[]
Factor 38	['second_connectivity_coh_freqmax_1']
Factor 39	['first_sp_abs_T']
Factor 40	['third_connectivity_coh_freqmax_2']
Factor 41	['first_ISI_Median']
Factor 42	[]
Factor 43	[]
Factor 44	['third_sp_abs_A']
Factor 45	[]

Factor	Features
Factor 46	['first_spectral_diff_1']
Factor 47	['second_amplitude_skew_2']
Factor 48	[]
Factor 49	[]
Factor 50	['second_spectral_diff_2']
Factor 51	['second_spectral_diff_1']
Factor 52	['second_IBI_length_median']
Factor 53	[]
Factor 54	['first_IBI_length_max']
Factor 55	[]
Factor 56	[]
Factor 57	[]
Factor 58	[]
Factor 59	[]
Factor 60	[]
Factor 61	[]

Table D.1: Factor Loadings for aEEG-EEG

Factor	Features
Factor 0	['volume - Cortical gray matter', 'volume - Anterior temporal lobe medial part right GM', 'volume - Superior temporal gyrus middle part right GM', 'volume - Occipital lobe right GM', 'volume - Occipital lobe left GM', 'volume - Frontal lobe right GM', 'volume - Frontal lobe left GM', 'volume - Parietal lobe right GM', 'volume - Parietal lobe left GM', 'volume - Temporal lobe right GM (merged region)', 'volume - Superior temporal gyrus right GM (merged region)']
Factor 1	['volume - Cerebellum', 'volume - Cerebellum left', 'volume - Cerebellum right']

Factor 2	['volume - Gyri parahippocampalis et ambiens anterior part left GM', 'volume - Gyri parahippocampalis et ambiens anterior part right GM', 'volume - Gyri parahippocampalis et ambiens right GM (merged region)']
Factor 3	['volume - CSF', 'volume - CSF.1']
Factor 4	['volume - Deep Gray Matter', 'volume - Thalamus right high intensity part in T2', 'volume - Thalamus left high intensity part in T2']
Factor 5	['volume - Caudate nucleus right', 'volume - Caudate nucleus left']
Factor 6	['volume - Anterior temporal lobe medial part right WM', 'volume - Anterior temporal lobe lateral part right WM']
Factor 7	['volume - Lentiform Nucleus right', 'volume - Lentiform Nucleus left']
Factor 8	['volume - Gyri parahippocampalis et ambiens posterior part right WM', 'volume - Gyri parahippocampalis et ambiens right WM (merged region)']
Factor 9	['Kido_WMscore40', 'Kido_Globalscore40']
Factor 10	['volume - Cingulate gyrus anterior part left GM', 'volume - Cingulate gyrus left GM (merged region)']
Factor 11	['volume - Thalamus right low intensity part in T2', 'volume - Thalamus left low intensity part in T2']
Factor 12	['volume - Brainstem', 'volume - Brainstem spans the midline']
Factor 13	['volume - Medial and inferior temporal gyri posterior part right WM', 'volume - Medial and inferior temporal gyri right WM (merged region)']
Factor 14	['volume - Subthalamic nucleus right', 'volume - Subthalamic nucleus left']
Factor 15	['volume - Cingulate gyrus posterior part right GM', 'volume - Cingulate gyrus right GM (merged region)']
Factor 16	['volume - Insula right GM', 'volume - Insula left GM']

Factor 17	['volume - Gyri parahippocampalis et ambiens posterior part right GM']
Factor 18	['Kido_CBscore40']
Factor 19	['volume - Intra-cranial background']
Factor 20	['volume - Cingulate gyrus anterior part right WM']
Factor 21	['volume - Gyri parahippocampalis et ambiens posterior part left GM']
Factor 22	['volume - White matter']
Factor 23	['volume - Gyri parahippocampalis et ambiens posterior part left WM']
Factor 24	['volume - Medial and inferior temporal gyri anterior part right WM']
Factor 25	['volume - Anterior temporal lobe lateral part right GM']
Factor 26	['Kido_CB_volumereduction40']
Factor 27	['volume - Ventricles']
Factor 28	['volume - Cingulate gyrus posterior part left GM']
Factor 29	[]
Factor 30	['volume - Cingulate gyrus anterior part right GM']
Factor 31	['volume - Superior temporal gyrus posterior part right GM']
Factor 32	['volume - Cingulate gyrus left WM (merged region)']
Factor 33	['volume - Cingulate gyrus posterior part right WM']
Factor 34	[]
Factor 35	[]
Factor 36	[]
Factor 37	['volume']
Factor 38	[]
Factor 39	[]
Factor 40	[]
Factor 41	[]

Table D.2: Factor Loadings for MRI Unscaled

Factor	Features
Factor 0	['volume - Cortical gray matter', 'volume - Anterior temporal lobe medial part right GM', 'volume - Superior temporal gyrus middle part right GM', 'volume - Occipital lobe left GM', 'volume - Superior temporal gyrus posterior part right GM', 'volume - Frontal lobe right GM', 'volume - Parietal lobe right GM', 'volume - Parietal lobe left GM', 'volume - Temporal lobe right GM (merged region)', 'volume - Superior temporal gyrus right GM (merged region)']
Factor 1	['volume - Cerebellum', 'volume - Cerebellum left', 'volume - Cerebellum right']
Factor 2	['volume - CSF', 'volume - CSF.1']
Factor 3	['volume - Gyri parahippocampalis et ambiens anterior part left GM', 'volume - Gyri parahippocampalis et ambiens anterior part right GM', 'volume - Gyri parahippocampalis et ambiens right GM (merged region)']
Factor 4	['volume - Cingulate gyrus anterior part left GM', 'volume - Cingulate gyrus left GM (merged region)']
Factor 5	['volume - Deep Gray Matter', 'volume - Thalamus right high intensity part in T2', 'volume - Thalamus left high intensity part in T2']
Factor 6	['volume - Anterior temporal lobe medial part right WM', 'volume - Anterior temporal lobe lateral part right WM']
Factor 7	['volume - Caudate nucleus right', 'volume - Caudate nucleus left']
Factor 8	['volume - Lentiform Nucleus right', 'volume - Lentiform Nucleus left']
Factor 9	['volume - Thalamus right low intensity part in T2', 'volume - Thalamus left low intensity part in T2']
Factor 10	['Kido_WMscore40', 'Kido_Globalscore40']

Factor 11	['volume - Medial and inferior temporal gyri posterior part right WM', 'volume - Medial and inferior temporal gyri right WM (merged region)']
Factor 12	['volume - Gyri parahippocampalis et ambiens posterior part right WM', 'volume - Gyri parahippocampalis et ambiens right WM (merged region)']
Factor 13	['volume - Brainstem', 'volume - Brainstem spans the midline']
Factor 14	['volume - Insula right GM', 'volume - Insula left GM']
Factor 15	['volume - Cingulate gyrus posterior part right GM', 'volume - Cingulate gyrus right GM (merged region)']
Factor 16	['volume - Gyri parahippocampalis et ambiens posterior part right GM']
Factor 17	['volume - Anterior temporal lobe lateral part right GM']
Factor 18	['volume - Intra-cranial background']
Factor 19	[]
Factor 20	['volume - Cingulate gyrus anterior part right WM']
Factor 21	['Kido_CBscore40']
Factor 22	['volume - Cingulate gyrus posterior part right WM']
Factor 23	['volume - Gyri parahippocampalis et ambiens posterior part left GM']
Factor 24	['volume - Medial and inferior temporal gyri anterior part right WM']
Factor 25	['volume - Subthalamic nucleus left']
Factor 26	['volume - White matter']
Factor 27	['volume - Gyri parahippocampalis et ambiens posterior part left WM']
Factor 28	['volume - Ventricles']
Factor 29	['volume - Cingulate gyrus anterior part right GM']
Factor 30	['Kido_CB_volumereduction40']
Factor 31	['volume - Cingulate gyrus posterior part left GM']
Factor 32	['volume - Frontal lobe left GM']
Factor 33	['volume - Subthalamic nucleus right']

Factor 34	['volume - Occipital lobe right GM']
Factor 35	[]
Factor 36	['volume - Cingulate gyrus left WM (merged region)']
Factor 37	[]
Factor 38	['volume']
Factor 39	[]
Factor 40	[]
Factor 41	[]

Table D.3: Factor Loadings for MRI Scaled wrt Age

Factor	Features
Factor 0	['volume - Cerebellum', 'volume - Cerebellum left', 'volume - Cerebellum right']
Factor 1	['volume - Cortical gray matter', 'volume - Frontal lobe right GM', 'volume - Frontal lobe left GM']
Factor 2	['volume - Parietal lobe right GM', 'volume - Parietal lobe left GM']
Factor 3	['volume - CSF', 'volume - CSF.1']
Factor 4	['volume - Gyri parahippocampalis et ambiens anterior part left GM', 'volume - Gyri parahippocampalis et ambiens anterior part right GM']
Factor 5	['volume - Anterior temporal lobe lateral part right GM', 'volume - Temporal lobe right GM (merged region)']
Factor 6	['volume - Deep Gray Matter', 'volume - Thalamus right high intensity part in T2', 'volume - Thalamus left high intensity part in T2']
Factor 7	['volume - Cingulate gyrus anterior part left GM', 'volume - Cingulate gyrus left GM (merged region)']
Factor 8	['volume - Lentiform Nucleus right', 'volume - Lentiform Nucleus left']

Factor 9	['volume - Medial and inferior temporal gyri posterior part right WM', 'volume - Medial and inferior temporal gyri right WM (merged region)']
Factor 10	['volume - Occipital lobe right GM', 'volume - Occipital lobe left GM']
Factor 11	['volume - Caudate nucleus right', 'volume - Caudate nucleus left']
Factor 12	['volume - Cingulate gyrus anterior part right GM', 'volume - Cingulate gyrus anterior part right WM']
Factor 13	['volume - Insula right GM', 'volume - Insula left GM']
Factor 14	['volume - Gyri parahippocampalis et ambiens posterior part right GM', 'volume - Gyri parahippocampalis et ambiens right GM (merged region)']
Factor 15	['Kido_WMscore40', 'Kido_Globalscore40']
Factor 16	['volume - Gyri parahippocampalis et ambiens posterior part right WM', 'volume - Gyri parahippocampalis et ambiens right WM (merged region)']
Factor 17	['volume - Superior temporal gyrus middle part right GM', 'volume - Superior temporal gyrus right GM (merged region)']
Factor 18	['volume - Brainstem', 'volume - Brainstem spans the midline']
Factor 19	['volume - Thalamus right low intensity part in T2', 'volume - Thalamus left low intensity part in T2']
Factor 20	['volume - Superior temporal gyrus posterior part right GM']
Factor 21	['volume - Gyri parahippocampalis et ambiens posterior part left GM', 'volume - Gyri parahippocampalis et ambiens posterior part left WM']
Factor 22	['volume - Cingulate gyrus posterior part right GM', 'volume - Cingulate gyrus right GM (merged region)']
Factor 23	['volume - Subthalamic nucleus right']
Factor 24	['volume - Medial and inferior temporal gyri anterior part right WM']
Factor 25	['volume - Cingulate gyrus posterior part left GM']

Factor 26	['volume - Anterior temporal lobe medial part right WM']
Factor 27	['volume - Intra-cranial background']
Factor 28	['volume']
Factor 29	['Kido_CBscore40']
Factor 30	['volume - Cingulate gyrus left WM (merged region)']
Factor 31	['volume - Cingulate gyrus posterior part right WM']
Factor 32	['volume - Ventricles']
Factor 33	['Kido_CB_volumereduction40']
Factor 34	['volume - Anterior temporal lobe medial part right GM']
Factor 35	['volume - Anterior temporal lobe lateral part right WM']
Factor 36	[]
Factor 37	['volume - White matter']
Factor 38	['volume - Subthalamic nucleus left']
Factor 39	[]
Factor 40	[]
Factor 41	[]

Table D.4: Factor Loadings for MRI Scaled wrt Age, Relative to Total Volume

Factor	Features
Factor 0	['volume - Cerebellum', 'volume - Cerebellum left', 'volume - Cerebellum right']
Factor 1	['volume - Cortical gray matter', 'volume - Frontal lobe right GM', 'volume - Frontal lobe left GM']
Factor 2	['volume - CSF', 'volume - CSF.1']
Factor 3	['volume - Deep Gray Matter', 'volume - Thalamus right high intensity part in T2', 'volume - Thalamus left high intensity part in T2']
Factor 4	['volume', 'volume - Ventricles']
Factor 5	['volume - Lentiform Nucleus right', 'volume - Lentiform Nucleus left']

Factor 6	['volume - Cingulate gyrus anterior part left GM', 'volume - Cingulate gyrus left GM (merged region)']
Factor 7	['volume - Parietal lobe right GM', 'volume - Parietal lobe left GM']
Factor 8	['volume - Gyri parahippocampalis et ambiens anterior part left GM', 'volume - Gyri parahippocampalis et ambiens anterior part right GM']
Factor 9	['volume - Medial and inferior temporal gyri posterior part right WM', 'volume - Medial and inferior temporal gyri right WM (merged region)']
Factor 10	['volume - Superior temporal gyrus middle part right GM', 'volume - Superior temporal gyrus right GM (merged region)']
Factor 11	['volume - Caudate nucleus right', 'volume - Caudate nucleus left']
Factor 12	['volume - Cingulate gyrus anterior part right GM', 'volume - Cingulate gyrus anterior part right WM']
Factor 13	['volume - Occipital lobe right GM', 'volume - Occipital lobe left GM']
Factor 14	['volume - Gyri parahippocampalis et ambiens posterior part right GM', 'volume - Gyri parahippocampalis et ambiens right GM (merged region)']
Factor 15	['volume - Insula right GM', 'volume - Insula left GM']
Factor 16	['Kido_WMscore40', 'Kido_Globalscore40']
Factor 17	['volume - Cingulate gyrus posterior part right GM', 'volume - Cingulate gyrus right GM (merged region)']
Factor 18	['volume - Gyri parahippocampalis et ambiens posterior part right WM', 'volume - Gyri parahippocampalis et ambiens right WM (merged region)']
Factor 19	['volume - Brainstem', 'volume - Brainstem spans the midline']
Factor 20	['volume - Thalamus right low intensity part in T2', 'volume - Thalamus left low intensity part in T2']
Factor 21	['volume - Superior temporal gyrus posterior part right GM']

Factor 22	['volume - Subthalamic nucleus right']
Factor 23	['volume - Medial and inferior temporal gyri anterior part right WM']
Factor 24	['volume - Gyri parahippocampalis et ambiens posterior part left GM', 'volume - Gyri parahippocampalis et ambiens posterior part left WM']
Factor 25	['volume - Anterior temporal lobe lateral part right GM']
Factor 26	['volume - Anterior temporal lobe medial part right WM']
Factor 27	['volume - Intra-cranial background']
Factor 28	['volume - Cingulate gyrus posterior part left GM']
Factor 29	['volume - Anterior temporal lobe medial part right GM']
Factor 30	['Kido_CBscore40']
Factor 31	['volume - Cingulate gyrus posterior part right WM']
Factor 32	['volume - Cingulate gyrus left WM (merged region)']
Factor 33	['volume - Anterior temporal lobe lateral part right WM']
Factor 34	['Kido_CB_volumereduction40']
Factor 35	['volume - White matter']
Factor 36	['volume - Temporal lobe right GM (merged region)']
Factor 37	[]
Factor 38	['volume - Subthalamic nucleus left']
Factor 39	[]
Factor 40	[]
Factor 41	[]

Table D.5: Factor Loadings for MRI Scaled wrt Age, Relative to (Total Volume - Ventricle Volume)

E. Appendix - MRMR Feature Selection

The following features were chosen by MRMR:

1. EEG: ['second_connectivity_coh_freqmax_1', 'second_sp_rel_B', 'first_-
   mse_slope_low', 'third_connectivity_coh_max_1', 'first_amplitude_skew_-
   4', 'second_spectral_diff_2', 'second_connectivity_coh_freqmax_3', 'first_-
   connectivity_coh_mean_2', 'third_amplitude_env_SD_4', 'third_rEEG_-
   width_3', 'first_FD', 'third_spectral_edge_frequency', 'first_spectral_-
   flatness_1', 'third_sp_abs_B', 'first_spectral_flatness_4', 'second_rEEG_-
   SD_3', 'third_rEEG_width_4', 'second_spectral_flatness_4', 'second_-
   amplitude_skew_1', 'first_rEEG_width_3', 'first_amplitude_skew_2',
   'first_connectivity_corr_3', 'first_rEEG_asymmetry_1', 'third_rEEG_-
   upper_margin_3', 'second_amplitude_skew_4', 'second_rEEG_asym-
   metry_4', 'first_amplitude_total_power_4', 'third_connectivity_BSI_-
   1', 'second_amplitude_SD_4', 'second_rEEG_median_4', 'second_asi_-
   hemi', 'first_rEEG_CV_1', 'third_amplitude_kurtosis_3', 'first_mse_-
   slope_high', 'third_connectivity_coh_freqmax_2', 'third_spectral_flat-
   ness_2', 'second_connectivity_coh_freqmax_2', 'third_amplitude_env_-
   mean_4', 'first_rEEG_upper_margin_3', 'third_amplitude_env_SD_1',
   'third_connectivity_coh_freqmax_1', 'first_spectral_entropy_3', 'first_-
   sp_abs_B', 'second_spectral_entropy_2', 'third_amplitude_kurtosis_-
   1', 'third_connectivity_BSI_2', 'second_spectral_diff_3', 'third_spectral_-
   flatness_4', 'third_spectral_diff_2', 'second_spectral_flatness_2', 'sec-
   ond_connectivity_BSI_2', 'first_spectral_entropy_4', 'first_amplitude_-
   env_SD_4', 'third_spectral_entropy_4', 'third_amplitude_total_power_-
   4', 'first_rEEG_SD_4', 'first_rEEG_mean_2', 'first_connectivity_coh_-
   freqmax_2', 'second_sp_abs_B', 'first_spectral_flatness_3', 'third_mse_-
   slope_high', 'second_spectral_diff_4']

2. MRI NS: ['Kido_WM_Myelinationdelay40', 'volume - Superior tem-
   poral gyrus middle part left WM', 'Kido_WM_Callosalthinning_MidMM40',

'volume - Thalamus right low intensity part in T2', 'volume - Superior temporal gyrus posterior part left GM', 'volume - Lateral Ventricle left', 'Kido_WM_VolumereductionMM40', 'volume - Cingulate gyrus anterior part left WM', 'Kido_GMscore40', 'Kido_DGMscore40', 'Kido_WM_cysticlesions40', 'volume - Insula right WM', 'volume - Brainstem', 'volume - Gyri parahippocampalis et ambiens anterior part right WM', 'Kido_GMscore_classes40', 'Kido_CB_signalabnormality40', 'Kido_WM_Callosalthinning40', 'volume - Cingulate gyrus posterior part left GM', 'volume - Cingulate gyrus anterior part right GM', 'volume - Anterior temporal lobe lateral part right WM', 'Kido_CB_volumereductionMM40', 'Kido_WM_Dilatedlateralventricles_leftMM40', 'volume - Anterior temporal lobe medial part left WM', 'Kido_WM_Dilatedlateralventricles40', 'volume - Anterior temporal lobe lateral part left WM', 'Kido_WM_Volumereduction40', 'Kido_GM_gyralmaturation40', 'volume - Background', 'Kido_DGM_signalabnormality40', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis anterior part right GM', 'Kido_WM_Callosalthinning_SpleniumMM40', 'volume - Gyri parahippocampalis et ambiens posterior part right WM', 'Kido_WM_focalsignalabnormality40', 'Kido_WM_Callosalthinning_GenuMM40', 'volume - Subthalamic nucleus left', 'volume - Cingulate gyrus right WM (merged region)', 'volume - Amygdala left', 'volume - Medial and inferior temporal gyri posterior part right WM', 'volume - Ventricles', 'volume - Medial and inferior temporal gyri anterior part right GM', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis left WM (merged region)', 'volume - Gyri parahippocampalis et ambiens posterior part right GM']

3. MRI VS: ['volume - Anterior temporal lobe lateral part left WM', 'volume - Background', 'Kido_WM_Dilatedlateralventricles_leftMM40', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis left WM (merged region)', 'Kido_DGMscore40', 'Kido_WM_Callosalthinning_MidMM40', 'volume - Lateral Ventricle left', 'volume - Ventricles', 'volume - Amygdala left', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis anterior part left WM', 'Kido_GMscore40', 'Kido_WM_Volumereduc-

tion40', 'volume - Cingulate gyrus anterior part left WM', 'volume - Brainstem', 'Kido_GMscore_classes40', 'Kido_WM_focalsignalabnormality40', 'volume - Medial and inferior temporal gyri posterior part right WM', 'volume - Insula left GM', 'Kido_WM_Callosalthinning_-SpleniumMM40', 'Kido_CBscore40', 'volume - Gyri parahippocampalis et ambiens posterior part right WM', 'Kido_GM_gyralmaturation40', 'Kido_WM_Callosalthinning_GenuMM40', 'Kido_WM_cysticlesions40', 'volume - Superior temporal gyrus posterior part right GM', 'volume - Cingulate gyrus anterior part right GM', ' age at scan', 'volume - Gyri parahippocampalis et ambiens posterior part left GM', 'volume - Cingulate gyrus right WM (merged region)', 'volume - Gyri parahippocampalis et ambiens anterior part right WM', 'volume - Subthalamic nucleus left', 'volume - Gyri parahippocampalis et ambiens posterior part right GM', 'Kido_WM_Callosalthinning40', 'Kido_DGM_-signalabnormality40', 'volume - Superior temporal gyrus middle part left WM', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis anterior part right WM', 'volume - Medial and inferior temporal gyri anterior part right GM', 'volume - Thalamus right low intensity part in T2', 'Kido_CB_signalabnormality40', 'volume - CSF', 'Kido_WM_Dilatedlateralventricles_rightMM40', 'volume - Anterior temporal lobe lateral part right WM']

4. MRI TVS: ['Kido_WM_Volumereduction40', 'Kido_WM_Callosalthinning_GenuMM40', 'volume - Superior temporal gyrus middle part left WM', 'volume - Superior temporal gyrus posterior part left GM', 'volume - Gyri parahippocampalis et ambiens posterior part right GM', 'volume - Medial and inferior temporal gyri anterior part left WM', 'Kido_CBscore40', 'volume - Background', 'volume - Thalamus right low intensity part in T2', 'volume - Insula right WM', ' age at scan', 'Kido_GM_increasedIHD40', 'volume - Medial and inferior temporal gyri anterior part right GM', 'volume - Medial and inferior temporal gyri anterior part right GM.1', 'volume - Thalamus left low intensity part in T2', 'volume - Gyri parahippocampalis et ambiens left GM (merged region)', 'Kido_WM_cysticlesions40', 'Kido_WM_Dilat-

edlateralventricles_rightMM40', 'volume - Gyri parahippocampalis et ambiens anterior part right WM', 'volume - Hippocampi and Amygdala', 'Kido_WM_VolumereductionMM40', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis posterior part left GM', 'volume - Superior temporal gyrus posterior part right WM', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis left WM (merged region)', 'volume', 'volume - Lateral Ventricle left', 'Kido_CB_signalabnormality40', 'Kido_GM_gyralmaturation40', 'volume - Insula left GM', 'volume - Subthalamic nucleus left', 'volume - Frontal lobe left WM', 'volume - Superior temporal gyrus posterior part left WM', 'Kido_DGMscore40', 'Kido_WM_Myelinationdelay40', 'volume - Ventricles', 'Kido_GMscore40', 'volume - Cerebellum', 'volume - Amygdala left', 'Kido_WM_focalsignalabnormality40', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis posterior part left WM', 'volume - Anterior temporal lobe lateral part right GM', 'volume - Cerebellum left']

5. MRI VVS: ['volume - Medial and inferior temporal gyri anterior part right GM.1', 'volume - Gyri parahippocampalis et ambiens left WM (merged region)', 'volume - Insula right GM', 'volume - Temporal lobe right WM (merged region)', 'Kido_CB_signalabnormality40', 'volume - Thalamus right low intensity part in T2', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis posterior part right GM', 'Kido_WM_focalsignalabnormality40', 'Kido_WM_Callosalthinning_MidMM40', 'volume - Brainstem', 'volume', 'volume - Gyri parahippocampalis et ambiens left GM (merged region)', ' age at scan', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis posterior part left GM', 'volume - Gyri parahippocampalis et ambiens posterior part left WM', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis anterior part right WM', 'Kido_WM_Volumereduction40', 'volume - Ventricles', 'Kido_GMscore40', 'volume - Medial and inferior temporal gyri anterior part right GM', 'Kido_WM_VolumereductionMM40', 'Kido_WM_Dilatedlateralventricles_rightMM40', 'Kido_DGMscore40', 'volume - Amygdala left', 'Kido_GM_gyralmaturation40', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis posterior part left WM', 'vol-

ume - Gyri parahippocampalis et ambiens anterior part left WM', 'volume - Insula left GM', 'volume - Subthalamic nucleus left', 'volume - Lateral Ventricle left', 'Kido_WM_cysticlesions40', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis posterior part right WM', 'volume - Hippocampi and Amygdala', 'Kido_CBscore_classes40', 'volume - Lateral occipitotemporal gyrus gyrus fusiformis left WM (merged region)', 'Kido_CBscore40', 'volume - Superior temporal gyrus posterior part right WM', 'volume - Gyri parahippocampalis et ambiens posterior part right GM', 'Kido_GM_increasedIHD40', 'volume - Insula right WM', 'volume - Thalamus left low intensity part in T2', 'volume - Superior temporal gyrus posterior part left GM']