EE 732
Probabilistic Graphical Models
Project Report
Mustafa Ümit ÖNER - 1675016
# Table of Contents

Table of Contents .................................................................................................................. 1
List of Figures .......................................................................................................................... 2
List of Tables ........................................................................................................................... 2
1. Introduction .......................................................................................................................... 3
   1.1. Functions of the Brain Sections ............................................................................... 3
2. Dataset and Data Quantization ......................................................................................... 4
   2.1. Data Quantization ........................................................................................................ 5
3. Structure Learning .............................................................................................................. 6
   3.1. Type1 DBN Structure .................................................................................................. 7
   3.2. Type2 DBN Structure ................................................................................................ 8
   3.3. Type3 DBN Structure ................................................................................................ 10
4. Analysis of the Structures ................................................................................................. 11
   4.1. Analysis over the Posterior Probability Matrices ...................................................... 11
      4.1.1. Analysis of Individual Subject Structures ......................................................... 12
      4.1.2. Analysis of 9-out-of-10 (K-fold) Technique Structures .................................... 14
   4.2. Analysis over the Connection of Nodes ..................................................................... 16
      4.2.1. Analysis of Individual Subject Structures ......................................................... 16
      4.2.2. Analysis of 9-out-of-10 (K-fold) Technique Structures .................................... 19
5. Conclusions ....................................................................................................................... 20
REFERENCES ......................................................................................................................... 22
Appendix .................................................................................................................................. 23
List of Figures

Figure 1: Position of the Electrodes on the Skull in EEG Recordings ........................................ 3
Figure 2: Lobes on Left View ........................................................................................................ 4
Figure 3: Lobes on Longitudinal Cross Section View .................................................................. 4
Figure 4: Functions on Left View ................................................................................................ 4
Figure 5: Functions on Longitudinal Cross Section View ............................................................ 4
Figure 6: Data Quantization Visual Representation ..................................................................... 5
Figure 7: Sample DBN Structure .................................................................................................. 6
Figure 8: Folder Indexing in storing Outputs of Structure Learning Algorithm ............................. 7
Figure 9: Sample Type1 DBN Structure ....................................................................................... 7
Figure 10: Sample Type2 DBN Structure ..................................................................................... 9
Figure 11: Sample Type3 DBN Structure ..................................................................................... 10
Figure 12: Yale Center for Dyslexia and Creativity - Reading Process Illustration ...................... 11
Figure 13: Connection Weights of Common Edges Shared by at least 5/10 of Control Before and Dyslexia Before Individuals ...................................................................................... 13
Figure 14: Connection Weights of Common Edges Shared by at least 5/10 of K-fold Control Before and Dyslexia Before Models ...................................................................................... 15
Figure 15: Comparison of Dyslexia Before Stimulus, and Control Before Stimulus Reachable Connections with Threshold Value of 0.2 by at least 5/10 of Individual Models.................... 17
Figure 16: Dyslexia Related Regions of the Brain ........................................................................ 18

List of Tables

Table 1: Type1 Posterior Probability Table ...................................................................................... 8
Table 2: Comparison of Mean Connection Distances with Threshold 0.2 and Lower Limit 5/10................................................................................................................................. 19
Table 3: Number of Reachable Common Connections with Threshold 0.2 and Lower Limit 5/10 .............................................................................................................................................. 19
Table 4: Comparison of Mean Connection Distances for K-fold Models with Threshold 0.2 and Lower Limit 10/10........................................................................................................................................... 20
Table 5: Number of Reachable Common Connections for K-fold Models with Threshold 0.9 .................................................................................................................................................. 20
Table 6: Comparison of Mean Connection Distances with Threshold 0.2 and Lower Limit 5/10................................................................................................................................. 21
1. Introduction

EEG recordings of subjects from two groups, group of subjects having dyslexia and control group of subjects who have not dyslexia, are given to be used in this project. These recordings are made on different regions of the subjects’ skulls, which is given with yellow shaded circles in the below figure.

![Figure 1: Position of the Electrodes on the Skull in EEG Recordings](image)

Aim of the project is to determine the effective connectivity of these regions by modeling this system as Dynamic Bayesian Network (DBN). Modelling will be done by using BNT and DBmcmc toolboxes.

In order to achieve the aim of the project, following tasks should be completed:

- Arrange the data
- Construct the DBN model
  - Learn the structure of the DBNs
  - Learn the parameters of the DBNs
- Analyze similarity of the constructed models

Before going detail of the analysis, it may be useful to have a look at main functional sections of the brain that the recordings are done and their functionalities.

1.1. Functions of the Brain Sections

Recordings are done at fur main functional lobes of the brain: frontal (F), parietal (P), temporal (T), and occipital (O). The region between frontal lobe and parietal lobe is called the Central (C) region. These regions are illustrated with their functions in Figure 2,3,4, and 5. The functions of these regions are declared in [1] as follows.
• **Frontal Lobe Functions:**
  - Attention and concentration
  - Self-monitoring
  - Organization
  - Speaking (expressive language)
  - Motor planning and initiation
  - Awareness of abilities and limitations
  - Personality
  - Mental flexibility
  - Inhibition of behavior
  - Emotions
  - Problem solving
  - Planning and anticipation
  - Judgment

• **Parietal Lobe Functions:**
  - Sense of touch
  - Spatial perception
  - Differentiation (identification) of size, shapes, and colors
  - Visual perception

• **Temporal Lobe Functions:**
  - Memory
  - Understanding language (receptive language)
  - Sequencing
  - Hearing
  - Organization

• **Occipital Lobe Functions:**
  - Vision

![Figure 2: Lobes on Left View](image)

![Figure 3: Lobes on Longitudinal Cross Section View](image)

![Figure 4: Functions on Left View](image)

![Figure 5: Functions on Longitudinal Cross Section View](image)

2. **Dataset and Data Quantization**

Data is given in two separate folders for dyslexia subjects and control group subjects. Each folder contains ten separate mat files one for each subject in each group. Each electrode’s recording is stored as separate matrices in mat files.

It is required to arrange these data for each subject as a single data matrix such that rows and columns correspond to electrodes and time points, respectively. After that, analog recordings must be quantized in order to be used with existing toolboxes. Some further processing is also
required since the data will be analyzed in two parts; before stimulus (which corresponds to
time slices of -500ms to -100ms) and after stimulus (which corresponds to time slices of
+100ms to +500ms).

2.1. Data Quantization

For each subject, each electrode’s measurements are horizontally concatenated (in the order of
O1; O2; P3; P4; P7; P8; T7; T8; C3; C4; F3; F4; F7; F8) to obtain each subject’s data matrix
for the test. After that, mean and standard deviation of each electrode’s measurements are
calculated and according to these values, measurements are quantized as -1 for the values less
than $\mu-\sigma$, +1 for the values greater than $\mu+\sigma$, and 0 for the values between $\mu-\sigma$ and $\mu+\sigma$, which
is also shown in Figure 6.

![Figure 6: Data Quantization Visual Representation](image)

After quantization is completed, before stimulus, and after stimulus data matrices for each
subject has been constructed from data points that correspondence to time slices of -500ms to
-100ms and 100ms to 500ms, respectively. Matlab script used for this process defined below.

Script name: data_quantization.m

Inputs : Analog electrode measurement mat files: d1,…,d10; k1,…,k10

Outputs : disleksi_before.mat
disleksi_after.mat
kontrol_before.mat
kontrol_after.mat

Parameters : None

Calling Fnc : None

Called Fnecs : None
3. Structure Learning

Structure learning will be conducted for each individual subject in each group, and for each group with 9-out-of-10 technique (this technique is referred as k-fold throughout the project), which keeps one subject out of analysis in each turn, both for before stimulus and after stimulus part of the data. As a result of structure learning process, a DBN, like in Figure 6, will be obtained.

![Figure 7: Sample DBN Structure](image)

Actually, three types of structures, i.e. three types of posterior probability matrices, each of which will be the different representations of the given data, will be constructed throughout the project, but the analysis will be conducted over the type3 structure. The structure types are described in the following sub-titles. Matlab script used for this process defined below.

Script name: structure_learning.m and structure_learning_kfold.m

Inputs:
- disleksi_before.mat
- disleksi_after.mat
- kontrol_before.mat
- kontrol_after.mat
- mcmcPAR.mat (Markov Chain Monte Carlo Simulation parameters)

Outputs:
- Posterior probability matrices and corresponding figures
- INTERposterior probability matrix

Parameters:
- threshold

Calling Fnc: None

Called Fnscs: DBmcmc_Application.m
- analysis.m

By running the structure learning script, first the DBmcmc_Application.m function is called and the function returns the INTERposterior probability matrix. Then, this matrix is given to the analysis.m function and three types of structures are obtained.

Structure learning script stores outputs in a folder indexing system as follows:
3.1. Type1 DBN Structure

In type1 structure, each electrode is treated individually and the corresponding model is constructed. Sample DBN structure and posterior probability matrix of this form is shown in Figure 8, and Table 1, respectively.

![Sample Type1 DBN Structure](image)
Table 1: Type1 Posterior Probability Table

<table>
<thead>
<tr>
<th>Time t+1</th>
<th>O1</th>
<th>O2</th>
<th>P3</th>
<th>P4</th>
<th>P7</th>
<th>P8</th>
<th>T7</th>
<th>T8</th>
<th>C3</th>
<th>C4</th>
<th>F3</th>
<th>F4</th>
<th>F7</th>
<th>F8</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>O2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0,1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P4</td>
<td>1</td>
<td>0,2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P7</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P8</td>
<td>0</td>
<td>0,1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>T7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0,4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>T8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>F3</td>
<td>1</td>
<td>1</td>
<td>0,1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0,9</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>F4</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0,9</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0,9</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>F7</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>F8</td>
<td>1</td>
<td>0</td>
<td>0,2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0,1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Script name: drawDBNType1.m
Inputs: type1_posterior
Outputs: Corresponding DBN structure
Parameters: threshold
Calling Fnc: analysis.m
Called Fncs: None

3.2. Type2 DBN Structure

For simplicity, electrodes are classified according to the hemisphere of the brain they exist and the ones that belong to the same lobe in a hemisphere are combined. Sample DBN structure and posterior probability matrix of this format is shown in Figure 9, and Table 2, respectively. In this structure, self-connections of the nodes are also omitted for simplicity. Squared nodes and circled nodes represent left and right hemisphere nodes on the skull, respectively.
Figure 10: Sample Type2 DBN Structure

Table 2: Type2 Posterior Probability Table

<table>
<thead>
<tr>
<th></th>
<th>OL</th>
<th>Or</th>
<th>Pl</th>
<th>Pr</th>
<th>Ti</th>
<th>Tr</th>
<th>Cl</th>
<th>Cr</th>
<th>Fl</th>
<th>Fr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time t</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Time t+1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0.1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Inputs: type2_posterior

Outputs: Corresponding DBN structure

Parameters: threshold

Calling Fnc: analysis.m

Called Fncts: None
3.3. Type3 DBN Structure

This representation is explicit illustration of Type2 DBN structure. Again, the left-right classification and node combination is done. Self-connections are also removed. Sample DBN structure looks like in Figure 10. Left nodes and right nodes are grouped at the left and the right of the figure, respectively. Sample posterior probability matrix is also given in Table 3.

![Sample DBN Structure](image)

**Figure 11: Sample Type3 DBN Structure**

**Table 3: Type3 Posterior Probability Table**

<table>
<thead>
<tr>
<th></th>
<th>O_l</th>
<th>P_l</th>
<th>T_l</th>
<th>C_l</th>
<th>F_l</th>
<th>O_r</th>
<th>P_r</th>
<th>T_r</th>
<th>C_r</th>
<th>F_r</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time t</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ol</strong></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><strong>Pl</strong></td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0,5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Tl</strong></td>
<td>1</td>
<td>1,4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Cl</strong></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Fl</strong></td>
<td>2</td>
<td>2,1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1,9</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Or</strong></td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0,1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Pr</strong></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0,3</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Tr</strong></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Cr</strong></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Fr</strong></td>
<td>1</td>
<td>2,2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0,9</td>
<td>1</td>
<td>1,1</td>
<td>4</td>
</tr>
</tbody>
</table>

- **Script name**: drawDBNType3.m
- **Inputs**: type2_posterior
- **Outputs**: Corresponding DBN structure
- **Parameters**: threshold
- **Calling Fnc**: analysis.m
- **Called Fnecs**: None
4. Analysis of the Structures

Analysis of the structures will be done over the type-3 posterior probability matrices of individual and k-fold models. Obtained matrices will be inputs to the analysis models.

It is declared by the Yale Center for Dyslexia and Creativity [2] about Dyslexia that:

“When you read, your brain has to do a lot of things at once. It has to connect letters with sounds and put those sounds together in the right order.

Then it has to help you put letters, words, and paragraphs together in ways that let you read them quickly and understand what they mean. It also has to connect words and sentences with other kinds of knowledge. When you see “c-a-t” on a piece of paper, your brain doesn’t just have to read the word “cat,” it also has to make the connection that “cat” means a furry, four-legged animal that meows.

Figure 12: Yale Center for Dyslexia and Creativity - Reading Process Illustration

Thanks to recent research, though, we have lots of scientific proof that a dyslexic person’s brain is normal and healthy.

When you have dyslexia, though, your brain takes longer to make some of these connections, and does it in more steps. It especially has trouble matching the letters you see on the page with the sounds those letters and combinations of letters make. And when you have trouble with that step, it makes all the other steps harder.”

According to this information, it is expected to see some similar connections between the subjects, especially for the control group. In order to see that, by thresholding the posterior probability tables, each connection will be tested for how many subjects have this connection. Over these values, common edges will be determined and analyzed. These calculations will be done for both types of the structures: individual structures and k-fold structures. This will constitute the first part of the analysis.

In the second part of the analysis, it is looked for the connections eventually set by the subjects based on the information that the dyslexia subjects set some of the connections in more steps. This will also be done for both structures, and this will be the second part of the analysis.

4.1. Analysis over the Posterior Probability Matrices

Posterior probability matrices (these are not exact probability matrices, rather proportional values to posteriors) are taken as input to the analysis. Type3 posterior probability matrix for
each structure, which is 10x10, is thresholded and converted to a row array of 1x100. Then, row arrays of each structure are vertically concatenated and 10x100 data matrices are obtained for the analysis.

Script name: data_arrangement.m and kfold_data_arrangement.m
Inputs: type3_posterior
Outputs: individual_data.xls and kfold_data.xls
Parameters: threshold
Calling Fnc: None
Called Fncs: None

4.1.1. Analysis of Individual Subject Structures

After obtaining data matrices, which are constructed from the posterior probability matrices of the individual subjects, each connection is searched for how many subjects’ structure contains this edge.

These values are used to determine an edge as common by setting a lower limit on these values. Edges with higher value than this limit, i.e. shared by more subjects than this limit value, are asserted as common edges. Common edges for each structure of individual subjects, dyslexia_before, dyslexia_after, control_before, and control_after, are determined, and cross comparison of common edges in each structure is done.

As an example analysis with threshold value of 0.2 and lower limit of 5, before stimulus type3 posterior probability matrices of dyslexia subjects are taken and thresholded with the value of 0.2 and the corresponding data matrix of dyslexia_before is obtained (data arrangement scripts are used). Then, each connection, which corresponds to the each column of the data matrix, is checked for how many subjects have this edge by counting non-zero rows in each column. Obtained numbers are compared with the lower value of 5 and the edges included at least in the 5/10 of the structures are considered as common.

Data matrix of control_before is also treated in a similar manner, and common edges for control after structures have been determined. Cross comparison of common edges in two before stimulus cases is done in a way that each common edge’s factor level and standard deviation is determined and plotted on the same figure to be analyzed. An example illustration is shown in Figure 13.

Similar, illustrations for all meaningful combinations is done: Control Before (CB) – Dyslexia Before (DB), Control After (CA) – Dyslexia After (DA), CB – CA, and DB – DA. They are all given in the project folder proposed in CD.
Figure 13: Connection Weights of Common Edges Shared by at least 5/10 of Control Before and Dyslexia Before Individuals
What is observed in the individual models is that the connections of individual models are divergent, i.e. they are not similar to each other at one glance.

4.1.2. Analysis of 9-out-of-10 (K-fold) Technique Structures

The k-fold posterior probability tables constructed for obtaining more reliable structures are used in this part of the analysis. The same procedure as in the previous part is followed.

Since k-fold models is obtained by concatenating the data of 9 models out of 10, the emergence of divergent connections of each individual subject with low connection weights in the structures is possible. In order to compensate that, the threshold value is increased to 0.9, which is reachable by common connections in all models. The connection weights of the two groups of control before and dyslexia before are shown in Figure 14.

During the experiment, common edges shared by 9/10 of the subjects of groups are not observed; however, at lower rates, such as 5/10, a lot of common edges exist between the two groups. This may results from the fact that threshold values for each subject is different, and some low evoke potential subjects may not contribute at sufficient strength to the k-fold models.

At this point, it is meaningful to analyze the ultimate reachable connections of models by using current available connections of the models.

Script name: individual_analysis.m and kfold_analysis.m

Inputs: type3_posterior

Outputs: DBNS_Shared… folders

Parameters: threshold

coeff % Lower limit on number of models that have connection

Calling Fnc: None

Called Fncs: None
Figure 14: Connection Weights of Common Edges Shared by at least 5/10 of K-fold Control Before and Dyslexia Before Models
4.2. **Analysis over the Connection of Nodes**

In this part of the analysis, type3 posterior probability tables of the models are used again. This time, the nodes that are reachable from a pre-determined node will be decided by using the breadth first searching algorithm, i.e. the connections related to this node will be found. By doing so, the directed distances of the connections will also be determined and comments related to whether a connection between two nodes is achieved ultimately, and the distance differences between the groups can be calculated.

Visual representations of the analysis similar to the previous part will be given and further comments will be done over those figures. In this case, mean values and standard deviations of distances of the connections will be used in the figures.

Similar to the previous part analysis, thresholded matrices are obtained. These matrices with one node as starting node are given as input to the breadth first search algorithm and distances of reachable nodes to this node are returned by the algorithm. These distances are stored as data matrices in excel files to be used in the following parts of the analysis.

```plaintext
Script name : connection_search.m and kfold_connection_search.m
Inputs      : type3_posterior
Outputs     : connectivity_distances.xls, kfold_connectivity_distances.xls
Parameters  : threshold
Calling Fnc : None
Called FnCs : bfs.m
```

```plaintext
Script name : bfs.m
Inputs      : adjacency matrix, start node
Outputs     : distances
Parameters  : None
Calling Fnc : connection_search and kfold_connection_search
Called FnCs : None
```

4.2.1. **Analysis of Individual Subject Structures**

In this part, distance data matrices are obtained by using type3 posterior probability tables of individual models. Again, the value that a connection is constructed in how many of the models is calculated, and it is asserted as common or not according to pre-determined lower limit value. Then, the related figures are constructed.

In Figure 15, cross comparison of dyslexia group before stimulus, and control group before stimulus reachable connections with threshold value of 0.2 by at least 5/10 of individual models is shown.
Figure 15: Comparison of Dyslexia Before Stimulus, and Control Before Stimulus Reachable Connections with Threshold Value of 0.2 by at least 5/10 of Individual Models
If there were a pathway in the brain related to the reading a word, it should be established in some way in all of the subject’s models. Therefore, it is meaningful to analyze the individual models in terms of ultimate connections and their distances. It should give more reliable evidences related to the dyslexia than the k-fold models, since k-fold models include all the connections of each individual model which makes k-fold models converges rapidly, i.e. set all the connections in short time, in some way.

In Table 2, and 3, the graphical illustration of Figure 15 is summarized in a compact manner. It can easily be seen that both dyslexia before and after models set more connections than control group models on the average, which is expected. Moreover, it is seen that nearly all the connections of the control before models are set by the dyslexia before models, which shows us that in the before stimulus region all subjects set the necessary connections (of common pathway), but dyslexia subjects do that in longer steps and time.

In [3], and [4], it is declared that developmental dyslexia is mainly related to the main three brain areas, which is also shown in Figure 16:

- Left parietotemporal region
- Left occipitotemporal region
- Left inferior frontal gyrus

![Figure 16: Dyslexia Related Regions of the Brain](image)

In Table 2, yellow shaded cells are related to the corresponding regions shown above and they are the connections that are set by dyslexia subjects in longer time than the control group subjects.]
Table 2: Comparison of Mean Connection Distances with Threshold 0.2 and Lower Limit 5/10

<table>
<thead>
<tr>
<th>Mean Connection Distances</th>
<th>Connections in CB but not in DB</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB &gt; CB</td>
<td>DB = CB</td>
</tr>
<tr>
<td>Cr*Cl</td>
<td>Pl*Pr</td>
</tr>
<tr>
<td>Cr*Or</td>
<td></td>
</tr>
<tr>
<td>Cr*Pr</td>
<td></td>
</tr>
<tr>
<td>Fl*Cl</td>
<td></td>
</tr>
<tr>
<td>Fl*Fr</td>
<td></td>
</tr>
<tr>
<td>Or*Ol</td>
<td></td>
</tr>
<tr>
<td>Pl*Cl</td>
<td></td>
</tr>
<tr>
<td>Pl*Fr</td>
<td></td>
</tr>
<tr>
<td>Pl*Ol</td>
<td></td>
</tr>
<tr>
<td>Pl*Or</td>
<td></td>
</tr>
<tr>
<td>Pr*Cr</td>
<td></td>
</tr>
<tr>
<td>Pr*Fl</td>
<td></td>
</tr>
<tr>
<td>Pr*Pl</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Number of Reachable Common Connections with Threshold 0.2 and Lower Limit 5/10

<table>
<thead>
<tr>
<th>Number of Reachable Common Connections</th>
<th>Number of Reachable Common Connections Between Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyslexia Before</td>
<td>61</td>
</tr>
<tr>
<td>Control Before</td>
<td>19</td>
</tr>
<tr>
<td>Dyslexia After</td>
<td>17</td>
</tr>
<tr>
<td>Control After</td>
<td>1</td>
</tr>
</tbody>
</table>

Script name: connectivity_analysis.m and kfold_connectivity_analysis.m

Inputs: connectivity distances

Outputs: distance figures

Parameters: dist_upper_limit % distance upper limit

Coeff %number of sharing models

Calling Fnc: None

Called Fncs: drawDBNType3.m

        drawCommonConn.m

4.2.2. Analysis of 9-out-of-10 (K-fold) Technique Structures

As it is declared in previous part of this report, it is meaningful to analyze the k-fold models constructed based on posterior probability matrices since they can give more information than the individual models constructed with limited data. However, it is preferable to look for individual models constructed based on connection distances since they show the ultimate connections in each individual model which is the focus of this study.
Although individual models are more reasonable than k-folds in this part of the analysis, Table 4, and 5 are constructed from k-fold models. When the number of connections in Table 5 is analyzed, it can easily be seen that nearly the entire after stimulus models reach each node, i.e. set each connection, and more than half of the before stimulus models do so. This is mainly due to emergence of divergent connections of individual subjects in k-fold models.

### Table 4: Comparison of Mean Connection Distances for K-fold Models with Threshold 0.2 and Lower Limit 10/10

<table>
<thead>
<tr>
<th>Mean Common Connection Distances</th>
<th>DB &gt; CB</th>
<th>DB = CB</th>
<th>DB &lt; CB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fl*Cl</td>
<td>Fr*Cl</td>
<td>Fl*Fr</td>
<td></td>
</tr>
<tr>
<td>Fr*Tl</td>
<td>Fr*Pl</td>
<td>Fl*Pl</td>
<td></td>
</tr>
<tr>
<td>Pl*Fl</td>
<td>Fr*Tr</td>
<td>Fl*Fl</td>
<td></td>
</tr>
<tr>
<td>Pl*Fr</td>
<td>Tr*Fl</td>
<td>Fr*Pl</td>
<td></td>
</tr>
<tr>
<td>Pl*Tr</td>
<td>Tr*Fl</td>
<td>Fl*Tl</td>
<td></td>
</tr>
<tr>
<td>Fl*Fr</td>
<td>Tr*Pl</td>
<td>Tl*Fr</td>
<td></td>
</tr>
<tr>
<td>Fl*Tl</td>
<td>Tr*Pl</td>
<td>Tl*Tl</td>
<td></td>
</tr>
<tr>
<td>Tl*Fl</td>
<td>Tr*Pl</td>
<td>Tl*Fl</td>
<td></td>
</tr>
<tr>
<td>Tl*Tr</td>
<td>Tr*Pl</td>
<td>Tl*Fr</td>
<td></td>
</tr>
</tbody>
</table>

### Table 5: Number of Reachable Common Connections for K-fold Models with Threshold 0.9

<table>
<thead>
<tr>
<th>Number of Reachable Common Connections Between Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Limit 10/10 9/10 8/10 5/10 10/10 9/10 8/10 5/10</td>
</tr>
<tr>
<td>Dyslexia Before 25 57 78 90 25 51 70 90</td>
</tr>
<tr>
<td>Control Before 56 71 82 90</td>
</tr>
<tr>
<td>Dyslexia After 81 90 90 90 81 90 90 90</td>
</tr>
<tr>
<td>Control After 90 90 90 90</td>
</tr>
</tbody>
</table>

### 5. Conclusions

Throughout the analysis, it is seen that modelling the pathways of dyslexia and control group subjects from their EEG recordings as Dynamic Bayesian Network gives consistent results with the literature in related field. It is shown that dyslexia subjects construct the same connections with the control group subjects in a longer time and using longer paths. Moreover, it is figured out that the common path related to reading a word is mostly related to the left hemisphere of the brain.

Analysis of the constructed structures are conducted mainly on two basis, one is based on posterior probability matrices, and the other one is the distances of the connections, i.e. how many steps it takes to reach a node, setting a connection. Actually, second part is also derived from posterior probability matrices, but it requires further computations. It is concluded that it is meaningful to analyze k-fold models in the first part since it compensates the scarcity of sample data. On the other hand, in the second part of the analysis, individual models give
more reasonable and meaningful results. When the important results of two parts, k-fold models in the first part, and individual models in the second part, are compared, they are also consistent with each other. Second part results table is given again to show the similarities of findings of two parts. Shared connections find in k-fold model structure analysis based on posterior probability matrices and individual model analysis based on distances are shaded in red in Table 6.

Table 6: Comparison of Mean Connection Distances with Threshold 0.2 and Lower Limit 5/10

<table>
<thead>
<tr>
<th>Mean Common Connection Distances</th>
<th>Connections in CB but not in DB</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB &gt; CB</td>
<td>DB = CB</td>
</tr>
<tr>
<td>Cr*Cl</td>
<td>Pr*Pr</td>
</tr>
<tr>
<td>Cr*Or</td>
<td>Pr*Ol</td>
</tr>
<tr>
<td>Cr*Pr</td>
<td>Pr*Fr</td>
</tr>
<tr>
<td>Fl*Cl</td>
<td>Pr*Cl</td>
</tr>
<tr>
<td>Fl*Fr</td>
<td></td>
</tr>
<tr>
<td>Or*Ol</td>
<td></td>
</tr>
<tr>
<td>Pl*Cl</td>
<td></td>
</tr>
<tr>
<td>Pl*Fr</td>
<td></td>
</tr>
<tr>
<td>Pl*Ol</td>
<td></td>
</tr>
<tr>
<td>Pl*Or</td>
<td></td>
</tr>
<tr>
<td>Pr*Cr</td>
<td></td>
</tr>
<tr>
<td>Pr*Fl</td>
<td></td>
</tr>
<tr>
<td>Pr*Pl</td>
<td></td>
</tr>
</tbody>
</table>

This study is a basis for further analysis. Since this is a directed model of the data flow in the brain, it can provide researchers with more information than the simple regional potential analysis of the brain. This is one of the strengths of this model over others in the literature. Moreover, it may leads to constructing artificial vision elements imitating the human vision system.
REFERENCES


Appendix

All the source code and library files are presented in the CD.

*Library files may be changed, so it is recommended to use given library files while re-run the given codes.*