Predicting Psychosis using the Experience Sampling Method with Mobile Apps

Presenter: Wajdi Alghamdi

ANDREA KATRINECZ, DANIEL STAMATE, WAJDI ALGHAMDI
DATA SCIENCE & SOFT COMPUTING LAB
GOLDSMITHS, UNIVERSITY OF LONDON, UK

PHILIPPE DELESPAUL
DEPARTMENT OF PSYCHIATRY AND PSYCHOLOGY
MAASTRICHT UNIVERSITY MEDICAL CENTRE, THE NETHERLANDS

DANIEL STAHL
DEPARTMENT OF BIOSTATISTICS AND HEALTH INFORMATICS
INSTITUTE OF PSYCHIATRY, PSYCHOLOGY & NEUROSCIENCE
KING’S COLLEGE LONDON, UK

JIM VAN OS
DEPARTMENT OF PSYCHIATRY, BRAIN CENTRE RUDOLF MAGNUS UNIVERSITY MEDICAL CENTRE UTRECHT, UTRECHT, THE NETHERLANDS
DEPARTMENT OF PSYCHIATRY AND PSYCHOLOGY
MAASTRICHT UNIVERSITY MEDICAL CENTRE, THE NETHERLANDS
KING’S HEALTH PARTNERS, DEPARTMENT OF PSYCHOSIS STUDIES
INSTITUTE OF PSYCHIATRY, PSYCHOLOGY & NEUROSCIENCE
KING’S COLLEGE LONDON, UK

SINAN GÜLOKSÜZ
DEPARTMENT OF PSYCHIATRY AND PSYCHOLOGY
MAASTRICHT UNIVERSITY MEDICAL CENTRE, THE NETHERLANDS
DEPARTMENT OF PSYCHIATRY
YALE SCHOOL OF MEDICINE, NEW HAVEN, CT, USA

16th IEEE International Conference on Machine Learning and Applications
ESM - Introduction

• *self-reporting* research method

• study people’s daily life during a *representative (typical) week*

• at *random moments* within every 2 hour blocks during waking hours stop and answer a set of questions > a snapshot of the mental state is obtained

• *unexpected* beeps to ensure natural behavior

• experience recorded in *real-time*

• **Question types:**
  • *Subjective situation*: thoughts, emotional, cognitive and motivational state – answer on a psychometric scale, typically *Likert scale* (level of agreement or disagreement along a range)
  • participants receive a *review* at the end
Advantages and Limitations

Introduction

- Advantages of ESM:
  - rich **longitudinal data** allows investigating dynamic flow of mental changes
  - **more accurate data**:
    - repeated assessments reduce assessment error
    - device is **always** with the participants
  - **cost effective**
    ▶ quicker and simpler process > improved participation > **larger sample sets**

- Limitations of ESM:
  - self-selection bias
  - social desirability bias
  - ESM procedure itself
Challenges in Psychiatry - Introduction

• Complex Assessment
  • Measurements of *psychological*, *biological* and *social factors*
  • Gathered from *interviews*, *examinations* and *medical history*

• Difficult Classification
  • *No clear boundaries* between classes
  • *Same symptoms* can indicate *different disorders*

• Treatment selection
  • *Individuals* respond differently

• Prediction of treatment outcome
  • Reducing treatment dosage can cause *relapse*
Objective

- **Objective**: distinguish patients from control using data collected by Experience Sampling Method (ESM) through Mobile Apps

- **Methodology**: Machine learning techniques such as data aggregation, data preprocessing, dimensionality reduction, prediction models, and Monte Carlo simulations.

- The **first** research to use ESM data via mobile applications to predict psychosis
Samples - Methodology

• pooled ESM-MERGE dataset:
  • 510 variables and 98,480 observations
  • Collected through the PsyMate mobile application

• Variables:
  • **Outcome variable:** status (categorical)
  • **Subjective predictor variables:** anxious, down, guilty, insecure, irritated, lonely, suspicious, cheerful, relaxed, satisfied > **Likert scale** (1-7)
  • **Demographic predictor variables:** age (continuous numeric) and sex (categorical)
  • **Other variables:** subject number, day number and beep number to help the aggregation process
Preprocessing - Methodology

• Only *psychotic patients* and *controls* kept.
• Only variables that were used in all studies were kept.
• Observations for only the first 6 days kept.
• *Correlation matrix* (Spearman): fairly high (0.24-0.67).
• *Missing values* (25% incomplete cases).

Dataset retained:
• 472 individuals (260 *patients* with psychosis and 212 *controls*).
• 60 observations from each individual.
Data Aggregation - Methodology

• Able to capture the variance in emotion
• Velocity, acceleration and abs(acceleration) calculated for each emotion variable

Sets created from:
• Base = Base data
• Velo = Base data + velocity
• Acc = Base data + velocity + acceleration
• Acc_abs = Base data + velocity + acceleration in absolute value
Data Aggregation - Methodology

• Replacing the 60 beeps of each individual with statistics, +add age and sex
  • Rule 1
    • Minimum
    • Maximum
    • 0.25 quantile
    • 0.5 quantile
    • 0.75 quantile
    • interquartile range
  • Rule 2
    • 0.1 quantile
    • 0.5 quantile
    • 0.9 quantile
    • interquartile range
Dimensionality Reduction - Methodology

A
• High correlation removal (e.g. Logistic Regression)
• Feature selection (Logistic Regression, Neural Nets, SVM)
  • Ranking by variable importance on Learning Vector Quantisation (LVQ) model
  • Recursive Feature Elimination (RFE)
  • ReliefF

B
• Principal Component Analysis
Machine Learning Algorithms - Methodology

• Logistic Regression
  • With and without StepAIC (stepwise model selection by Akaike Information Criterion)

• Support Vector Machines
  • With linear, polynomial and radial kernels

• Gaussian Processes
  • With linear, polynomial and radial kernels

• Neural Networks
  • With single hidden layer

• Random Forest
  • 2000 trees
Model Training and Tuning - Methodology

• Nested cross-validation
  • 5-fold outer cross validation
  • 10-fold inner cross validation with ROC used to find the best probability cut-offs

• Monte Carlo simulation
  • 5 repetitions to approximate whether the model performs well
  • 100 repetition run on well performing models to test stability
Hardware and Software - Methodology

• Computationally expensive procedure > robust framework needed

• Hardware Infrastructure
  • Data analytics cluster with parallel processing
  • 11 servers with Intel Xeon processors
  • 832GB fast RAM

• Software
  • R software
  • Packages: caret, pROC, MASS, e1071, CORElearn, randomForest, ggplot2, data.table, mclust, stringi, spatstat, plyr, DMwR, arm, AppliedPredictiveModeling, doParallel and kernlab
Predictive Modelling - Results

• Best 20 models:
  • Aggregated by Rule 2
  • Include base, velocity and acceleration in normal values
  • Relief feature selection
  • Algorithms:
    • SVM with radial and polynomial kernel
    • Gaussian Process with radial kernel
    • Random Forest
  • PCA

• Top 3:
  • SVM with radial kernel (82% Accuracy, 82% Sensitivity)
  • SVM with polynomial kernel (80% Accuracy, 79% Sensitivity)
  • Gaussian Process with radial kernel on PCA set (79% Accuracy, 78% Sensitivity)
## Predictive Modelling - Results

### TABLE II. BEST RESULTS OF THE 3 TOP PERFORMING MODELS

<table>
<thead>
<tr>
<th>Method</th>
<th>Dataset</th>
<th>ROC</th>
<th>Sens</th>
<th>Spec</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM Radial</td>
<td>Rule 2</td>
<td>0.8639</td>
<td>0.8192</td>
<td>0.8255</td>
<td>0.8220</td>
<td>0.6419</td>
</tr>
<tr>
<td>SVM Polynomial</td>
<td>Rule 2</td>
<td>0.8435</td>
<td>0.7885</td>
<td>0.8113</td>
<td>0.7987</td>
<td>0.5959</td>
</tr>
<tr>
<td>GP Radial</td>
<td>Rule 2 PCA</td>
<td>0.8216</td>
<td>0.7808</td>
<td>0.7925</td>
<td>0.7860</td>
<td>0.5700</td>
</tr>
</tbody>
</table>

### TABLE III. MONTE CARLO 100 EXPERIMENTS
AVERAGE RESULTS OF THE 3 TOP PERFORMING MODELS

<table>
<thead>
<tr>
<th>Method</th>
<th>Dataset</th>
<th>ROC</th>
<th>Sens</th>
<th>Spec</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM Radial</td>
<td>Rule 2</td>
<td>0.8459</td>
<td>0.7706</td>
<td>0.7957</td>
<td>0.7819</td>
<td>0.5623</td>
</tr>
<tr>
<td>SVM Polynomial</td>
<td>Rule 2</td>
<td>0.8300</td>
<td>0.7481</td>
<td>0.7828</td>
<td>0.7637</td>
<td>0.5264</td>
</tr>
<tr>
<td>GP Radial</td>
<td>Rule 2 PCA</td>
<td>0.8157</td>
<td>0.7582</td>
<td>0.7535</td>
<td>0.7561</td>
<td>0.5093</td>
</tr>
</tbody>
</table>
Predictive Modelling - Results

Accuracy

Sensitivity

Specificity

Kappa
Feature Analysis - Results

- Dataset including velocity and acceleration, aggregated by Rule 2
- The three feature selection results were compared
- Both velocity and acceleration were useful addition, with more acceleration showing amongst the top 10
- Most representative measures:
  - 0.9 quantile, 0.1 quantile, interquartile range
  - Anxious, insecure, suspicious in both emotional level and level change form
  - Cheerful, feeling down and lonely in emotional level form
Conclusion

• ML methods detected predictive patterns
• Results consistent with researches: variance in emotion changes is beneficial in predicting patients
• Further research into multilevel structure
• Building a detection system for mental illnesses
• Include also clinical, behavioural, genetic, environmental variables
• Apply advanced techniques such as deep learning.
Questions?