Mathematical Institute

Using a signature-based machine learning model to analyse a psychiatric stream of data

Imanol Perez
(Joint work with T. Lyons, K. Saunders and G. Goodwin)

Mathematical Institute
University of Oxford

Rough Paths in Toulouse

Oxford Mathematics

- Given $p \geq 1$ and $X \in \mathcal{C}([s,t], \mathbb{R}^d)$ with $s < t$ we define

$$\|X\|_{p,[s,t]} := \sup_{\{t_i\}_i \subset [s,t]} \left(\sum_i \|X_{t_i} - X_{t_{i-1}}\|^p \right)^{1/p}.$$

- Given $p \geq 1$ and $X \in \mathcal{C}([s,t], \mathbb{R}^d)$ with $s < t$ we define

$$\|X\|_{p,[s,t]} := \sup_{\{t_i\}_i \subset [s,t]} \left(\sum_i \|X_{t_i} - X_{t_{i-1}}\|^p \right)^{1/p}.$$

- $\mathcal{V}^p([s,t], \mathbb{R}^d) := \{X \in \mathcal{C}([s,t], \mathbb{R}^d) : \|X\|_{p,[s,t]} < \infty\}$.

Definition (Signature of a continuous path)

Let $X \in \mathcal{V}^1([0,T],\mathbb{R}^d)$. The signature of X is defined as

$$S(X) = (1, X^1, X^2, \ldots) \in \bigoplus_{n=0}^{\infty} (\mathbb{R}^d)^{\otimes n}$$

where

$$X^n = \int \ldots \int_{0<u_1<u_2<\ldots<u_n<T} dX_{u_1} \otimes \ldots \otimes dX_{u_n} \quad \forall n \geq 1.$$

Definition (Truncated signature of a continuous path)

Similarly, we define, for $n \geq 0$,

$$S^n(X) := (1, X^1, X^2, \ldots, X^n).$$

Definition (Time-joined transformation)

Let $\{(t_i, X_{t_i})\}_{i=0}^{N} \subset \mathbb{R}^+ \times \mathbb{R}$ be a stream of data. Its time-joined transformation is defined as the path $Y : [0, 2N+1] \to \mathbb{R}^+ \times \mathbb{R}$ that is given by

$$Y_t := \begin{cases} (t_0, X_{t_0} t) & \text{for } t \in [0,1) \\ (t_i + (t_{i+1} - t_i)(t - 2i - 1), X_{t_i}) & \text{for } t \in [2i+1, 2i+2) \\ (t_{i+1}, X_{t_i} + (X_{i+1} - X_{t_i})(t - 2i - 2)) & \text{for } t \in [2i+2, 2i+3) \end{cases},$$

for $0 \leq i \leq N-1$.

Signature of a stream of data

Mathematical Institute

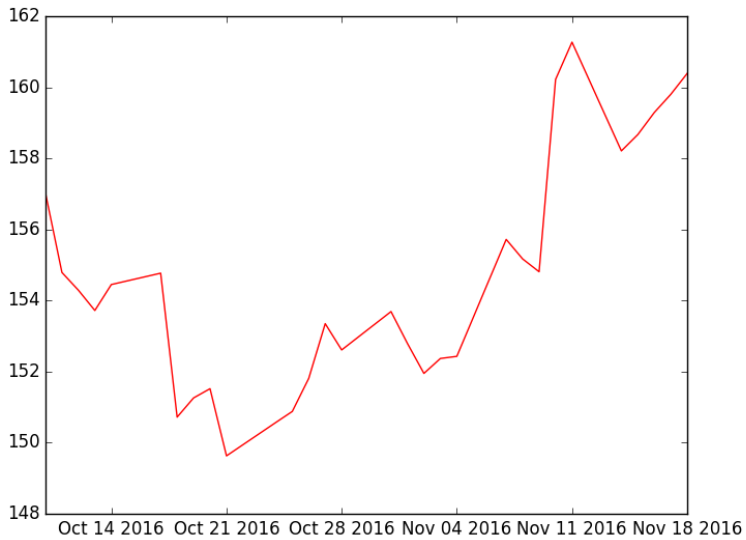


Figure: IBM stock price from October to November 2016.

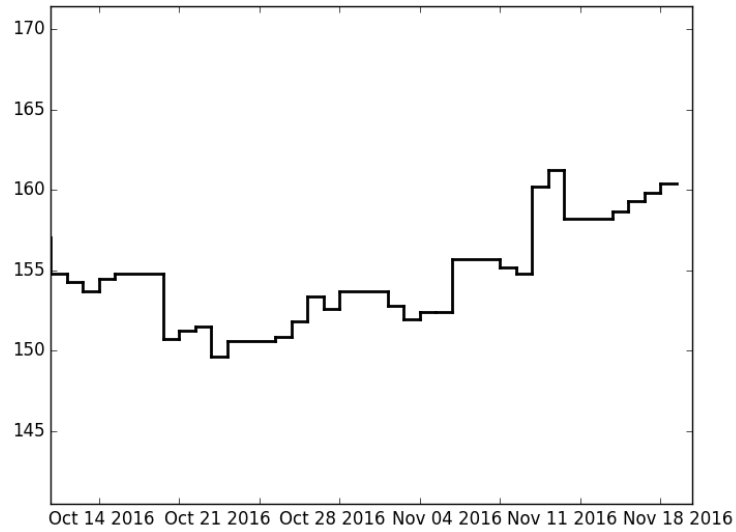


Figure: Time-joined transformation of the path.

Definition (Signature of a stream of data)

The signature of a stream of data $\{(t_i, X_{t_i})\}_{i=0}^{N}$, which with some abuse of notation will be denoted by $S(\{(t_i, X_{t_i})\}_{i=0}^{N})$, is defined as the signature of its time-joined transformation.

- We have two data sets: a known set of known input-output pairs (the *training set*), $\{(X_i, Y_i)\}_i$, which is used to train the model, and a set of inputs that is used for testing (the *out-of-sample set*).

- We have two data sets: a known set of known input-output pairs (the *training set*), $\{(X_i, Y_i)\}_i$, which is used to train the model, and a set of inputs that is used for testing (the *out-of-sample set*).
- Features play an important role in machine learning.

UNIVERSITY OF OXFORD
Mathematical Institute

Theorem (B. Hambly, T. Lyons)

The signature of a path with bounded variation is unique up to tree-like equivalence.

Signatures and machine learning

Signatures as features: estimate

Theorem

Let $X \in \mathcal{V}^1([0,T],\mathbb{R}^d)$ be a path with bounded variation. Then, given $1 \leq i_1, i_2, \ldots, i_n \leq d$ we have

$$\left\| \int \ldots \int_{0<u_1<u_2<\ldots<u_n<T} dX_{u_1}^{i_1} \ldots dX_{u_n}^{i_n} \right\| \leq \frac{\|X\|_{1,[0,T]}^n}{n!}.$$

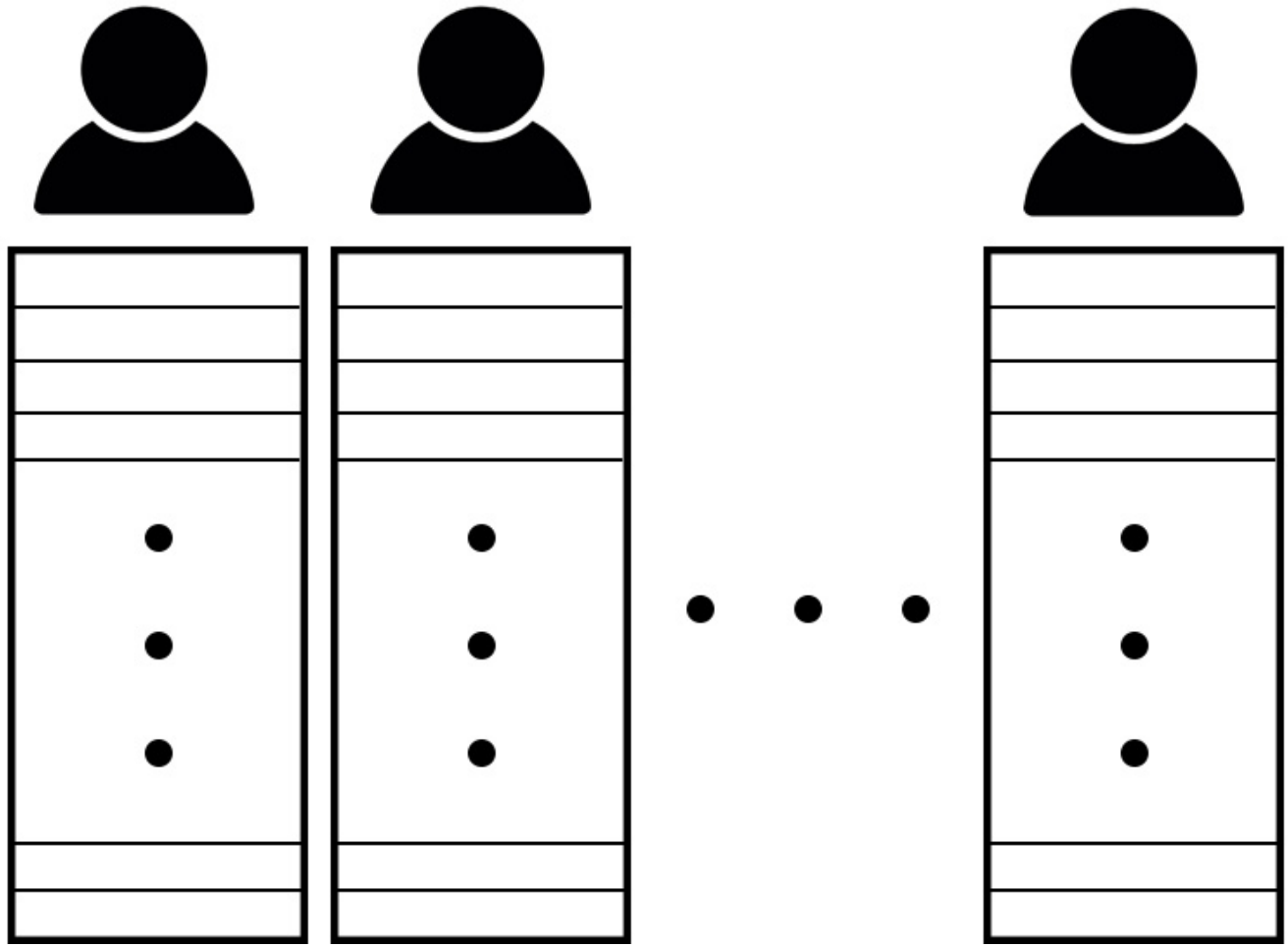
- Given a training set $\{(R_i, Y_i)\}_{i=0}^N$, of input-output pairs, where $R_i = \{(t_{ij}, r_{ij})\}_j$ is a stream of data, construct a new set $\{(X_i, Y_i)\}_{i=0}^N$ with $X_i \in \mathcal{V}^1$.

- Given a training set $\{(R_i, Y_i)\}_{i=0}^N$, of input-output pairs, where $R_i = \{(t_{ij}, r_{ij})\}_j$ is a stream of data, construct a new set $\{(X_i, Y_i)\}_{i=0}^N$ with $X_i \in \mathcal{V}^1$.
- Compute $\{(S^n(X_i), Y_i)\}_{i=0}^N$ for some $n \in \mathbb{N}$.

- Given a training set $\{(R_i, Y_i)\}_{i=0}^N$, of input-output pairs, where $R_i = \{(t_{ij}, r_{ij})\}_j$ is a stream of data, construct a new set $\{(X_i, Y_i)\}_{i=0}^N$ with $X_i \in \mathcal{V}^1$.
- Compute $\{(S^n(X_i), Y_i)\}_{i=0}^N$ for some $n \in \mathbb{N}$.
- Apply regression against the truncated signature.

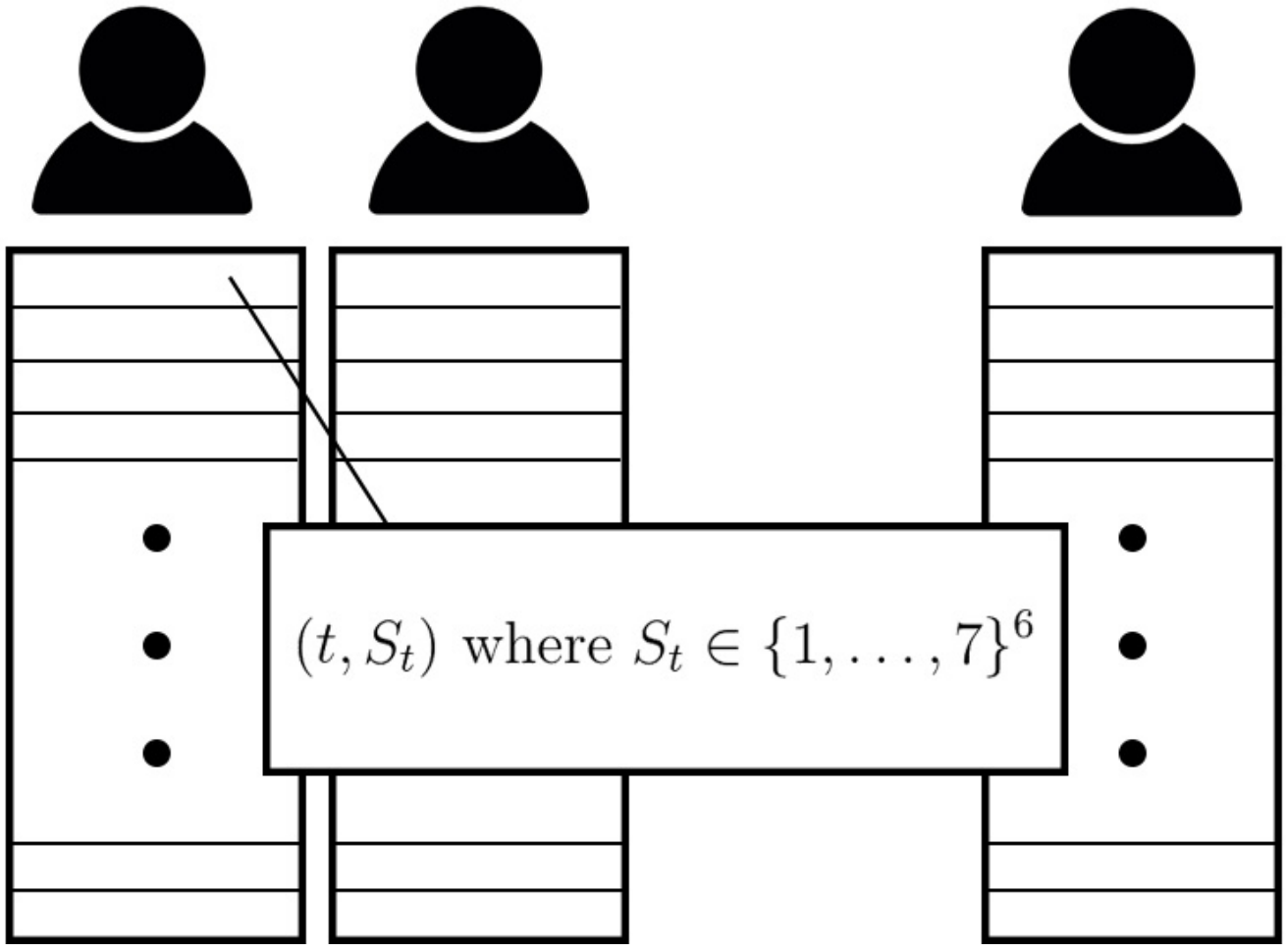
Application to psychiatric data

The problem



Application to psychiatric data

The problem

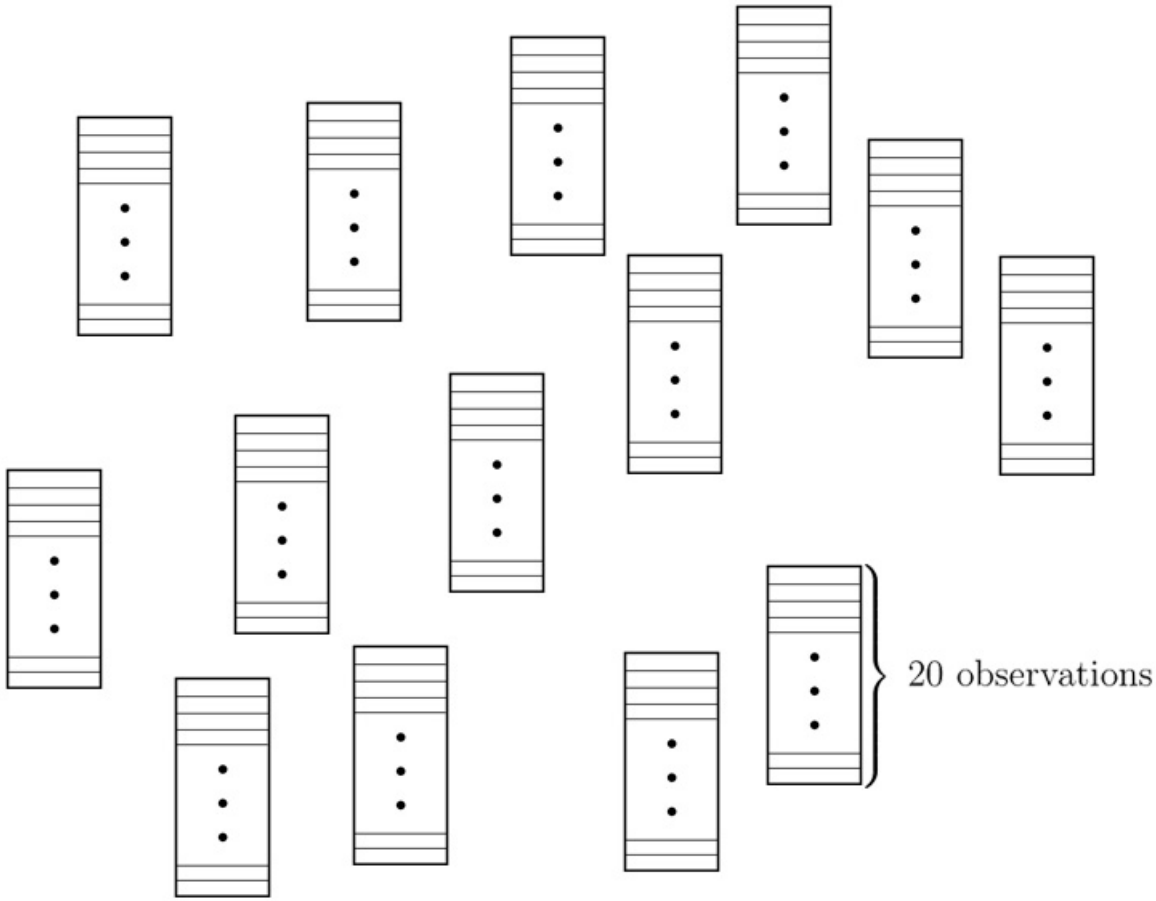


- Given some information about a participant, can we tell if he or she was diagnosed to have bipolar disorder, borderline personality disorder or to be healthy?

- Given some information about a participant, can we tell if he or she was diagnosed to have bipolar disorder, borderline personality disorder or to be healthy?
- Given a participant and information about the last few days, can we predict the mood the following day?

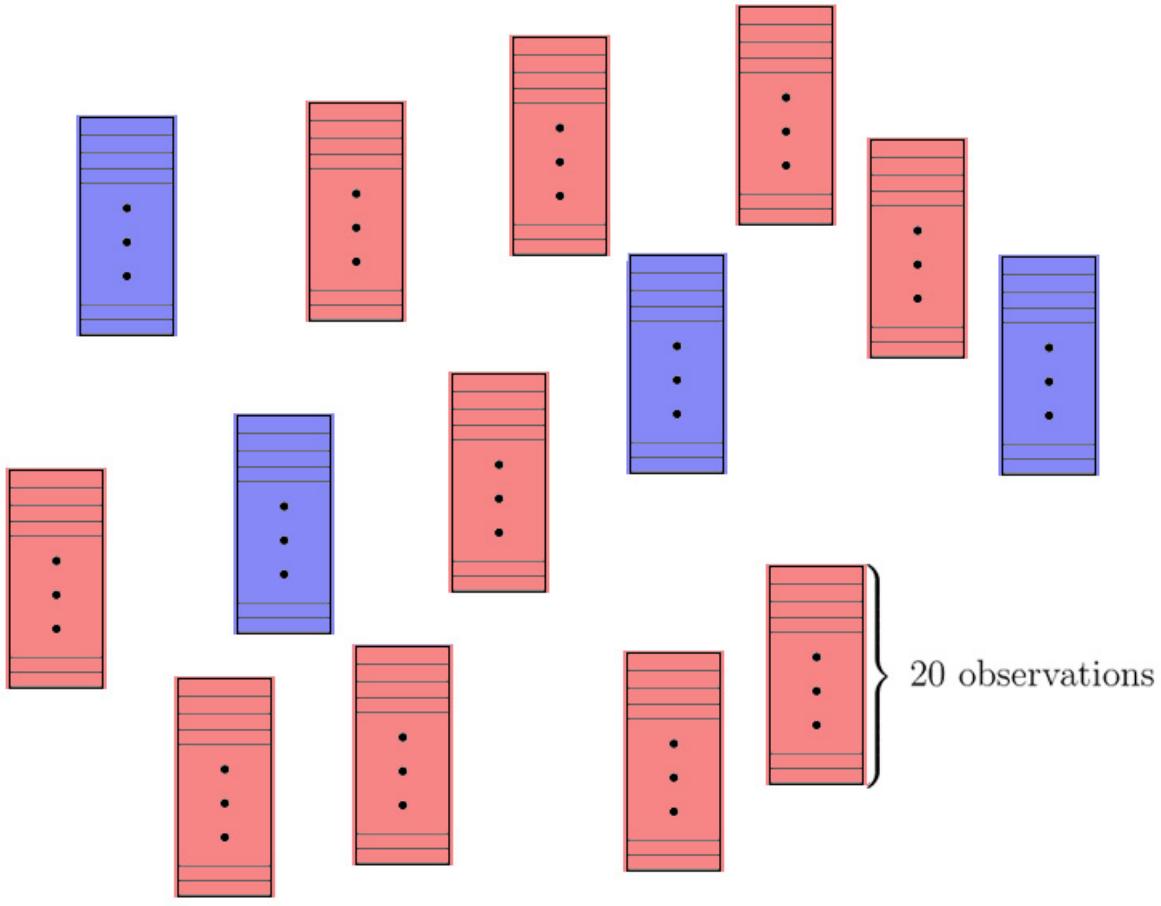
Application to psychiatric data

Methodology



Application to psychiatric data

Methodology



Application to psychiatric data

Methodology

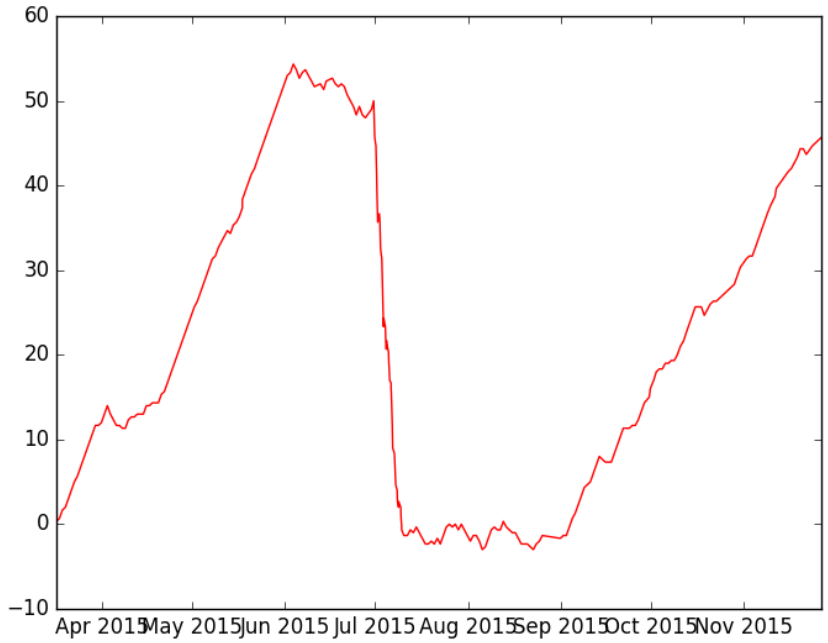
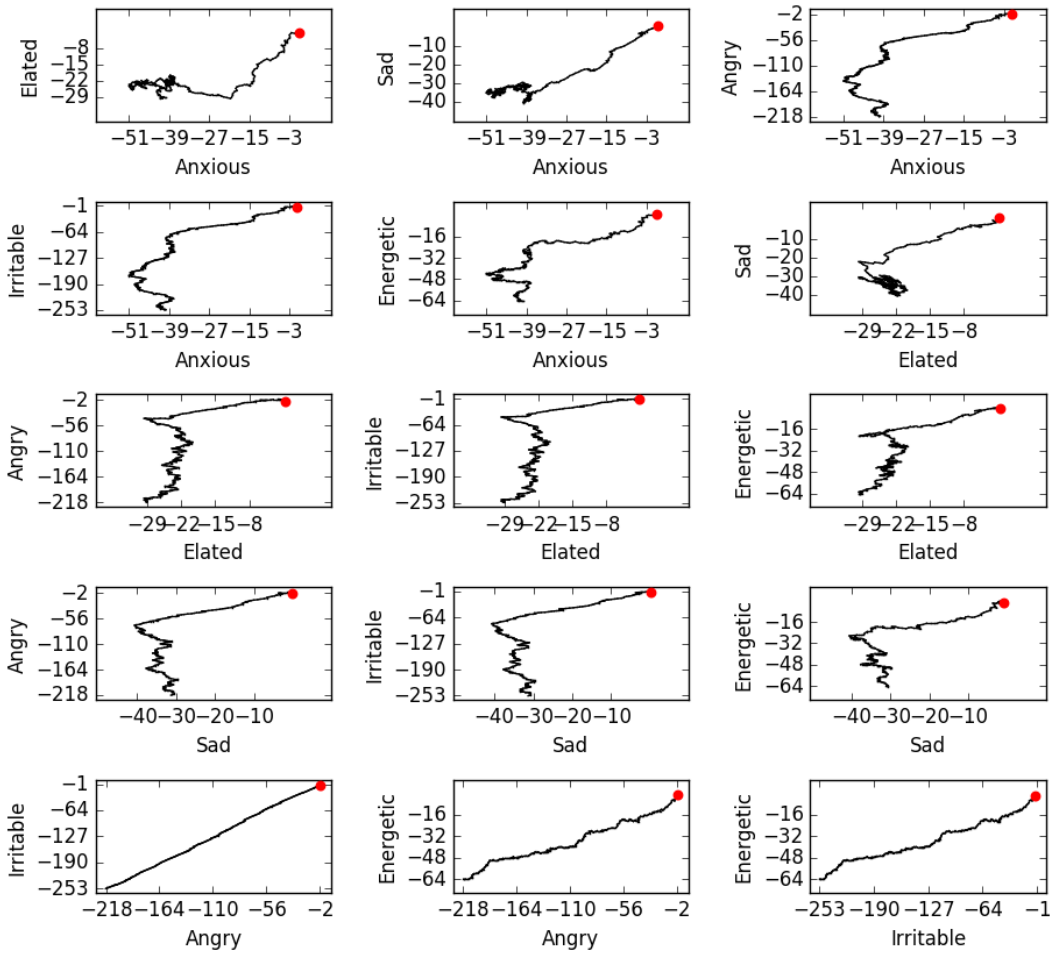


Figure: Normalised path for anxiety scores.

Application to psychiatric data

Methodology



Application to psychiatric data

Predicting if a person is healthy, has bipolar disorder or has borderline disorder

UNIVERSITY OF OXFORD
Mathematical Institute

$$\{(t_i, S_{t_i})\}_{i=0}^{19} \quad \rightarrow \quad \begin{cases} (-1,1), & \text{if the partcipant is healthy} \\ (-1,-1), & \text{if the participant is bipolar.} \\ (1,0), & \text{if the participant is borderline.} \end{cases}$$

Application to psychiatric data

Predicting if a person is healthy, has bipolar disorder or has borderline disorder

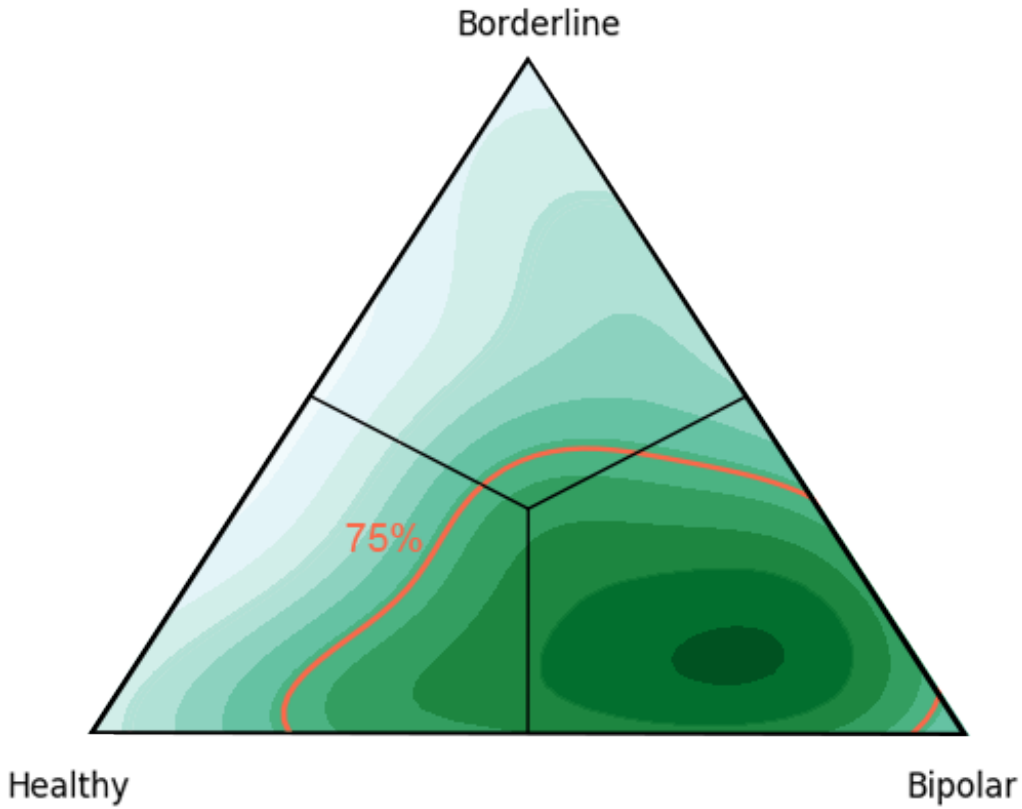
Order	Correct guesses
2nd	75%
3rd	70%
4th	69%

Table: Percentage of people correctly classified in the three clinical groups.

Application to psychiatric data

Predicting if a person is healthy, has bipolar disorder or has borderline disorder

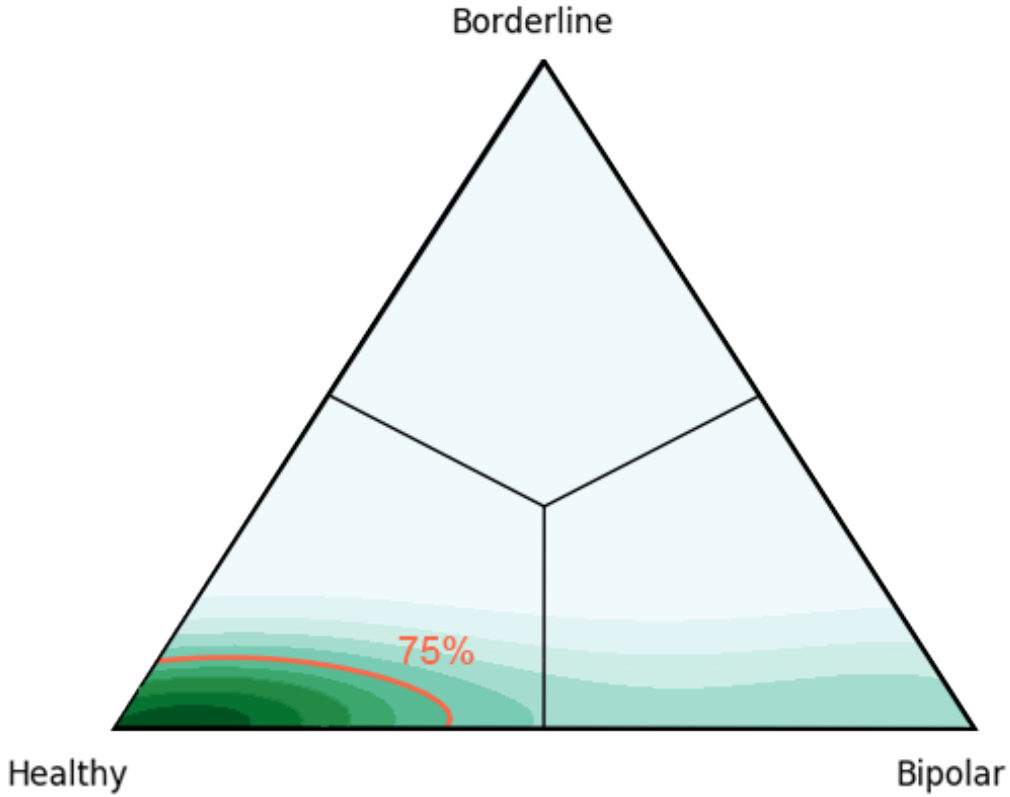
UNIVERSITY OF OXFORD
Mathematical Institute



Application to psychiatric data

Predicting if a person is healthy, has bipolar disorder or has borderline disorder

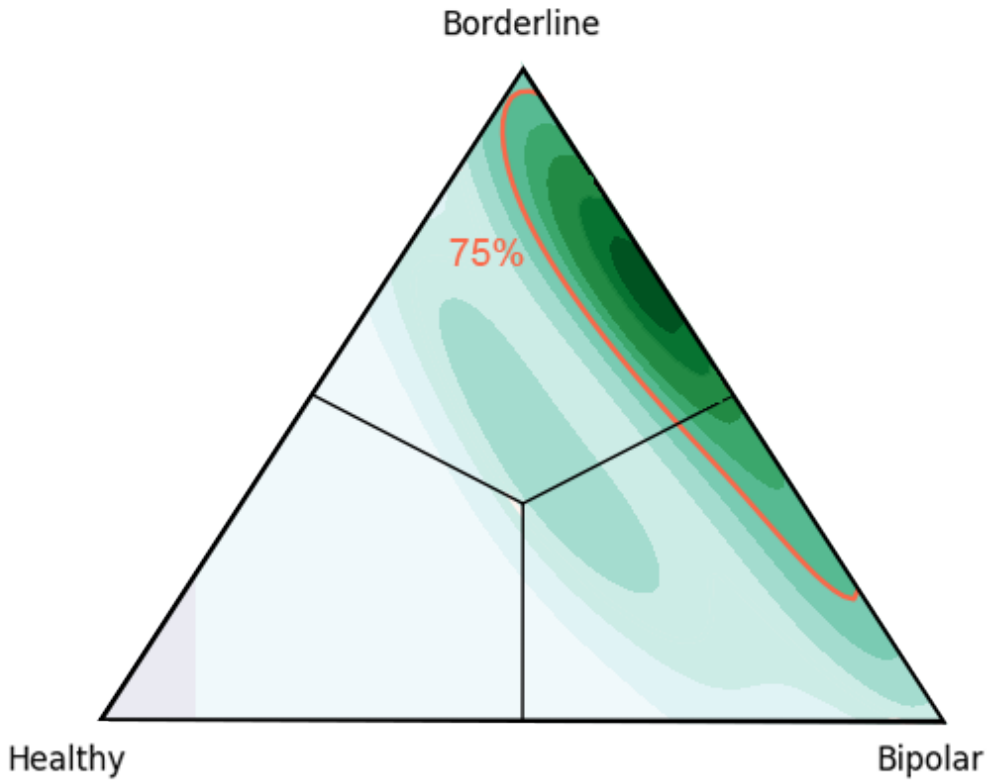
UNIVERSITY OF OXFORD
Mathematical Institute



Application to psychiatric data

Predicting if a person is healthy, has bipolar disorder or has borderline disorder

UNIVERSITY OF OXFORD
Mathematical Institute



$$\{(t_i, S_{t_i})\}_{i=0}^{19} \quad \rightarrow \quad S \in \{1, \ldots, 7\}^6$$

where S is the scores of the participant the following observation.

Application to psychiatric data

Predicting the future mood

UNIVERSITY OF OXFORD
Mathematical Institute

Category	Healthy	Bipolar	Borderline
Anxious	98%	82%	73%
Elated	89%	86%	78%
Sad	93%	84%	70%
Angry	98%	90%	70%
Irritable	97%	84%	70%
Energetic	89%	82%	75%

Table: Percentage of correct guesses for mood predictions

Thank you!