

# Deep learning for bipolar disorder: anticipating manic and depressive episodes

Shreya Shankar (shreya@cs.stanford.edu)

Stanford University, Department of Computer Science

## Predicting

- Bipolar disorder, a mental illness, is defined as a brain disorder that causes unusual shifts in mood, energy shifts, and activity levels
- A person with bipolar disorder shifts between manic (high) and depressive (low) episodes
- Most people frequently engage in text messaging as one of the more dominant forms of communication in their lives
- Is it possible to predict in advance when I will have manic and depressive episodes based on my text messages?
- Built LSTM autoencoder to identify anomalies in text message data and correlate them with episodes

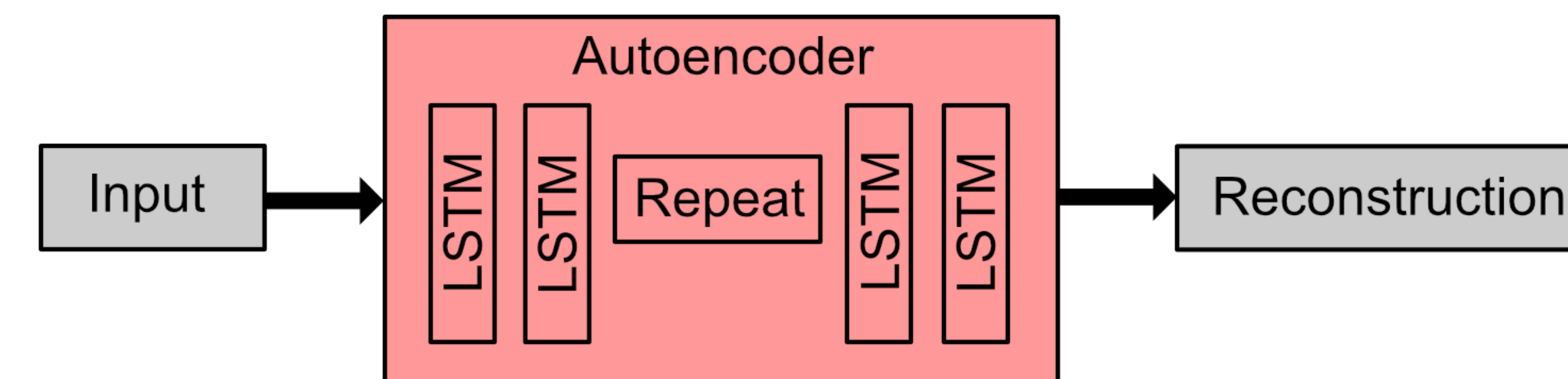
## Data

- 60,000 text messages I sent from March 8, 2019 until October 15, 2019 from iMessage and Messenger
- 100-dimensional GloVe embeddings trained on text data from Twitter [1]
- Stripped punctuation and lowercased all the words, in addition to removing stop words
- 7 weeks labeled with hypomanic or depressive episodes, but labels not used during training model

## Features

- Text messages split into words, where each word is converted into a vector
- Each text message is associated with a timestamp
- Messages capped at 50 words: shorter messages are zero-padded, while longer messages are truncated
- GloVe vectors [1] are fixed and not updated during training

## Model



- LSTM autoencoder to minimize mean squared error: encoder has two LSTM layers, decoder has two LSTM layers
- $f$  represents LSTM function

$$\mathbf{h}_0 = \mathbf{0}$$

$$\mathbf{h}_t = f(\mathbf{W}\mathbf{h}_{t-1} + \mathbf{U}\mathbf{x}_t + \mathbf{b})$$

$$\mathbf{y}_t = \text{sigmoid}(\mathbf{W}_0\mathbf{h}_t + \mathbf{b}_0)$$

$$J(\mathbf{y}) = \frac{1}{n} \sum_{i=1}^n ((\mathbf{y}_t)_i - (\mathbf{x}_t)_i)^2$$

## Training results



## Results

- A data point counts towards precision if it lies within one of the weeks labeled by the psychiatrist data
- A week in the psychiatrist data is "recalled" if at least one of the model's predicted dates lies in that week.
- Set MSE threshold to 500,000
- Daily reconstruction errors above threshold are considered part of "episodes"

Data split	Average MSE	Size of split	Precision	Recall
Train	325340.48	36,000	64%	57%
Dev	373148.62	12,000	56%	43%
Test	384620.92	12,000	57%	43%

## Discussion

- Model overfits somewhat to the training set
- Model did well in reconstructing short text, poorly in reconstructing long texts that seemed to be copy-pasted from other sources (i.e. medicine description, error message found while debugging, etc)
- Overall, did not expect results to be better than 50%, since text message data is hard to analyze (lots of shorthand, etc)
- Text data contents probably not as indicative of episodes as text message frequency, so it would be nice to include that

## Future

- Construct a k-Means clustering baseline on the frequencies of text message, or even use text message frequency as a feature
- Separate predicting mania and depression into two separate problems
- More rigorous hyperparameter tuning and methods to reduce overfitting

## References

- [1] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. Glove: Global vectors for word representation. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, 2014.