Can Machine learning be used to predict bipolar states with different types of input data?

Joakim I. Frogner

Thesis submitted for the degree of
Master in Programming and Networks
60 credits

Department of Informatics
Faculty of mathematics and natural sciences

UNIVERSITY OF OSLO

Spring 2018

Can Machine learning be used to predict bipolar states with different types of input data?

Joakim I. Frogner

Can Machine learning be used to predict bipolar states with different types of input data?

http://www.duo.uio.no/

Printed: Reprosentralen, University of Oslo

Abstract

Contents

List of Figures

List of Tables

Preface

Part I

Introduction

Chapter 1

Introduction

1.1 Motivation

Statistics
- Data shows that 5,890,000 adults are diagnosed with bipolar disorder in the USA (2,65% of the adult population) [find better source].

Ways to use the results of this study?

1.2 Thesis overview

[Fill in later]

Chapter 2

Background

2.1 Bipolar disorder

Bipolar disorder is the disorder where you experience extreme mood swings. One day you can feel amazing and everything is fine, but the next day you feel like you don't belong anywhere in this universe. Mood swings in general is not something that you should be concerned about. It is however the extreme cases where your mind turns 180 degrees from day to day that is the main symptom of bipolar disorder. There is not really a specific type of people that get this; they can be of any age and any gender, but most people that suffer from it find out (by having an experience or episode) around age 25 [3].

When talking about bipolar disorder, we often separate between the states *normal*, *mania* and *depression*. The last two are the states we usually talk about, since a normal state isn't that interesting. These two states are very different, but they have some similarities, for example sleeping problems.

When a bipolar person is in a manic state, he/she may do things that they never would have intended doing, like spending a lot of money on items they really don't need, or abusing drugs/alcohol. They may also feel really excited or powerful [2].

A bipolar patient is in a depressive state when he or she is in a bad mood swing. They can stop doing everything they usually like to do, and lie down in bed all day with no motivation to do anything useful. They may feel useless and that they don't belong here, or being guilty of something they may or may not have done. In some cases, a depression may even end up with suicidality, where the person either just thinks of death, or actually attempt suicide (actually 20% of people diagnosed with bipolarity commit suicide [3]).

The frequency of these symptoms can vary. One year they can have these mood swings every day for several weeks at the time, and the next they get them less frequent, like once every month.

We also separate between bipolar disorder type I and II, with the main difference being that the manic episodes are way more aggressive in type I [1].

Statistics say that bipolarity is genetically inheritable, with 23% chance of getting a child with bipolar disorder if one parent is bipolar, and 66% if both parents are [3].

2.2 Machine learning

Machine learning is the field of computer science where you basically throw a lot of data into an algorithm and expect it to give you answers to whatever you prefer, with as little work as possible. This was not the case in the early days of the technology, but nowadays it is a lot easier with all the diffent frameworks and tools available.

Machine learning is a great and almost 'magical' technology, but only if you do it right. First you need to have enough data to feed into the algorithm, and to be efficient when training the model on a large dataset, which you need to be if you want your result quickly, you need good hardware. You can get away with a decent CPU if you just want to test it out on a small dataset, but if you really want to do machine learning, then you need a good GPU. The reason why GPUs are so much better than CPUs on this specific task, is because the CPUs are designed for flexibility and general computing workloads. The GPUs on the other hand, are designed to do the same instructions over and over again in parallel. This makes GPUs a lot more efficient for machine learning, and especially for deep neural networks [5].

Now how do you do the actual machine learning? Well there are many diffent approaches to this, which I will discuss in the next sections, but say you want to use a neural network for your task. Then your next step should be to choose a framework. You can of course do everything yourself, but why reinvent the wheel when there are so many good frameworks and tools already out there?

The programming language **Python** is great for machine learning in my (and many other peoples) opinion. It is structured in a way such that it really looks like pseudo-code, and this is perfect because we don't want to spend time on weird syntax rules in another language. For Python, you have a popular framework called **TensorFlow** which is developed by Google. This allows you to build models easily, and also execute the training and testing. Before you get started with TensorFlow, do a quick google search to see if someone else has already done something similiar to what you are trying to acheive, and if you find something, odds are that your neural network model can be similar, if not identical to it. If not, then you have to sit down and actually make the model yourself.

For the model implementation part, whether you found a model online or want to build it yourself, you can of course do it in TensorFlow, but there is an easier way. **Keras** is also a popular framework that is most commonly used together with TensorFlow. On their documentation website [4], they see their framework as 'A high-level neural networks API, written in Python and capable of running on top of TensorFlow, CNTK or Teano'.

Follwing their '30 seconds to Keras' guide [4], you can create a 'sequential' model with 'dense' layers, configure its learning process (compile), then fit, train, evaluate and predict with just a few lines of code:

Source Code 2.1: 30 Seconds to Keras

```
from keras.models import Sequential
from keras.layers import Dense

model = Sequential()

model.add(Dense(units=64, activation='relu', input_dim=100))
model.add(Dense(units=10, activation='softmax'))

model.compile(loss='categorical_crossentropy',
              optimizer='sgd',
              metrics=['accuracy'])

model.fit(x_train, y_train, epochs=5, batch_size=32)
loss_and_metrics = model.evaluate(x_test, y_test, batch_size=128)
classes = model.predict(x_test, batch_size=128)
```

So, as long as you know your theory, and can decide which model to use (and either find a good implementation of that model or create it yourself), you really just have to make the dataset ready. This is the boring and tedious part of machine learning, but it has to be done in order for making it possible to train the model on it.

2.3 Machine learning strategies

2.3.1 Supervised learning

2.3.2 Unsupervised learning

2.3.3 Semi-supervised learning

2.4 Machine learning approaches

2.4.1 Decision tree learning

2.4.2 Reinforcement learning

2.4.3 Neural networks

- General idea
- Deep learning

2.5 How can machine learning help

How I can solve the problem described in the beginning of this chapter...

2.6 Related work

List related work and discuss...

2.7 Challenges

Describe some challenges...

2.8 Ethical concerns

Describe some ethical concerns...

Part II

The project

Chapter 3

Planning the project

Part III

Conclusion

Chapter 4

Results

Bibliography

[1] *Bipolar 1 Disorder and Bipolar 2 Disorder: What Are the Differences?* Accessed: 14.03.2018. URL: https://www.healthline.com/health/bipolar-disorder/bipolar-1-vs-bipolar-2.

[2] *Bipolar disorder*. Accessed: 14.03.2018. URL: https://familydoctor.org/condition/bipolar-disorder/.

[3] *Bipolar disorder statistics*. Accessed: 14.03.2018. URL: https://www.statisticbrain.com/bipolar-disorder-statistics/.

[4] *Keras documentation*. Accessed: 15.03.2018. URL: https://keras.io.

[5] *Why are GPUs well-suited to deep learning?* Accessed: 15.03.2018. URL: https://www.quora.com/Why-are-GPUs-well-suited-to-deep-learning.