Using Convolutional Neural Networks To Analyze Social Media Trends For Stock Market Analysis

Mizzani Walker-Holmes, Nathanael Leman, Kushum Thapa, Mark Malaney, Isaac Efik
Computer Science Department
Kennesaw State University
Marietta, Georgia 30060 USA

Faculty Advisor: Dr. Meng Han

Abstract

The mission of the AI For Finance project is to provide tools to promote financial literacy through engagement and education. The objective of this project is to create tools to assist individuals in using data to influence their financial well-being regarding investment decisions made in the stock market. This project leverages sentiment analysis tooling, deep learning models, and big social data mining to analyze stock potential and maximize future returns based on company's social media presence, stock market data history and linear regression calculations. Social media analytics was collected from Twitter and Instagram to accompany stock market data in order to better gauge public sentiment on identified companies and forecast their stock potential. Pre-processed datasets were analyzed through a facial recognition pipeline using a convolutional neural network model in TensorFlow and vector embedding to detect patterns and classify images for further sentiment analysis. Findings and deliverables were summarized and collected to reside in a user-friendly web application called the "AI 4 finance" application. This is an interactive interface in which users can see the mechanisms behind the tools that we created and obtain better insight on the various sources of data that can be leveraged to conduct sound financial investments. Through the visualizations and information provided on the application, individuals can better tailor their financial game plans and portfolios regarding stock market forecasting and make fiscally responsible investment decisions into successful stock outcomes.

Keywords: Machine Learning, Convolutional Neural Networks, Computer Vision, Crowdsourcing, Pattern Recognition

1. Introduction

Regardless of background or income level, financially informed individuals are better able to take control of their circumstances, improve their quality of life, and ensure a stable future for themselves and their surroundings. This project seeks to inspire financial independence and increase individuals' long- term professional and personal growth.

   Our objective is to create tools to assist individuals in using data to influence their financial well-being regarding investment decisions made with in the stock market. This project leverages sentiment analysis tooling, deep learning models, and big social data mining to analyze stock potential and maximize future returns based on the company's social media presence, stock market data history and linear regression calculations.

2. Methods

In this research two methods of data collection were used for the stock market analysis Facial Recognition System: 1) social media data analytics analyzing the text and image data from Facebook, Twitter and Instagram and 2) data from the stock market was obtained by using an API provided by IEX Trading. Both sources of data collected were used to train the conventional neural network model in TensorFlow.

2.1. Social Media Data

Twitter data was retrieved using Twitter Developer API to analyze how companies present themselves in the public space through content such as advertisement campaigns and consumer satisfaction. Instagram data was also retrieved using the Github repo

"Instagram Scraper" to analyze images associated with the company's visual representation. supervised sentiment analysis tools were used to calculate and categorize the negative, positive and neutral social media data collected on identified companies.

2.2. Stock Market Data

To generate stock market data API calls were all made through a Python program that will query the API for all relevant stock data for the specified companies and store the data in a csv file or JSON file. Using this API, we were able to gather the daily opening and closing price, the 52-week high and low, the volume, and many other data endpoints. The IEX API allowed us to retrieve all of this data about all 100 companies that were tracked for a full five years of data per company.

3. Results

Data collected from social media and the stock market was processed and trained through a facial recognition pipeline with deep learning created in TensorFlow. In order to increase accuracy of the data collected, datasets were cleaned, trained and pre-processed to comprised inputs that included companies with viable market data, social media images and well-defined public sentiment. Trained datasets were then pushed through a convolutional neural network model using softmax loss with the Incepcetion-Resnet-v1 model as well as FaceNet that directly learns a mapping from face images to compact space where distance correspond to face measurements and similarities. This system allowed our model to directly detect detect patterns and classify images for further analysis on social media images and text collected on Facebook, Twitter and Instagram.

4. Conclusion

This research project successfully demonstrated that data science principles could be used to accurately predict stock market trajectory and further encourage positive financial decision making among users. Using diverse data resources from social media and the stock market individuals can be equipped to analyze the stock market potential of companies and make wise investments. A user-friendly interactive application was developed to allow users to select which companies they'd like to invest into and see based on the corresponding social media trends and stock market data available how likely it would be for that particular company to increase or decrease in stock market success. Users could also tailor the application around their financial goals, build custom portfolios and read financial tips for better stock market forecasting and financial planning.

5. Acknowledgements

6. References

1. Xiaoyan Shao, Siyuan Lu, Theodore G. van Kessel, Hendrik F. Hamann, Leda Daehler, Jeffrey Cwagenberg, Alan Li, *Solar irradiance forecasting by machine learning for solar car races*, Big Data (Big Data) 2016 IEEE International Conference on, pp. 2209-2216, 2016.

2. Kuen-Han Tsai, Yau-Shian Wang, Hsuan-Yu Kuo, Jui-Yi Tsai, Ching-Chih Chang, Hui-Ju Hung, Hong-Han Shuai, *Multi-Source Learning for Sales Prediction*, Technologies and Applications of Artificial Intelligence (TAAI) 2017 Conference on, pp. 148-153, 2017.