

Abstract

This thesis investigates the use of computational finance and machine learning tools for identifying investment opportunities in fashion retail assets. The thesis consists of three key experiments:

Experiment 1: Celebrity and brand influence on the popularity of the fashion news

The first experiment, which focuses on the collection and the basic analysis of the dataset, sets the background to the following chapters. More specifically, it investigates the correlation between the celebrities and brands mentioned on the fashion news and the popularity of the news on hypebeast.com and hypebae.com. The datasets for this experiment are collected by the self-developed crawler from hypebeast.com and hypebae.com, which are the most influential websites on men's fashion and culture. It contains 182,294 fashion news labeled by their categories, keywords which including the related celebrities and brands, and the index of popularity.

Experiment 2: Text analysis to predict the popularity of the fashion news

The second experiment use NLP to create a classifier to analyse the news data, predict the popularity, and derive metrics, which could be used as the input for further modeling. This experiment explores various algorithms for text classification for fashion news data. In this experiment, the dataset is performed basic pre-processing and spitted into train and validation sets. The feature engineering is also implemented which transforms the raw data into flat features and also create new features

from the existing data. The effectiveness of traditional models are evaluated and the best performing classifier is used in the final model in the following experiment.

Experiment 3:

The third experiment proposes a framework to predict the stock price movement using various data sources, including e-commerce data, social media, financial data, and identify investment opportunities. The framework is in the form of a [*the introduction to the final algorithm I use in the model. As I did not start to implement the final model, I will keep it for later.*]. The implementation of the model is made available in the Python programming language.

Contributions to Science

Due to lack of the understanding how to analyze the fashion retailing, the investors often overlook the niche financial securities such as stocks of the public fashion companies. The main research contribution of this thesis is to explore new methodology to investment decision making based on different online data sources, such as e-commerce, social network, financial economic ,and fashion news.

This thesis contributes to the existing literature in a number of ways. First, this research proposes a new view to analyze the fashion news and investigates the correlation between the celebrities and fashion trends. Second, the research examines the effectiveness of different algorithm for fashion news popularity prediction and derive the metric for the stock movement prediction model. Third, the final framework is designed combining previous text analysis and other public online data sources, which are capable to undertake pre-investment analysis and recommend investment opportunities.