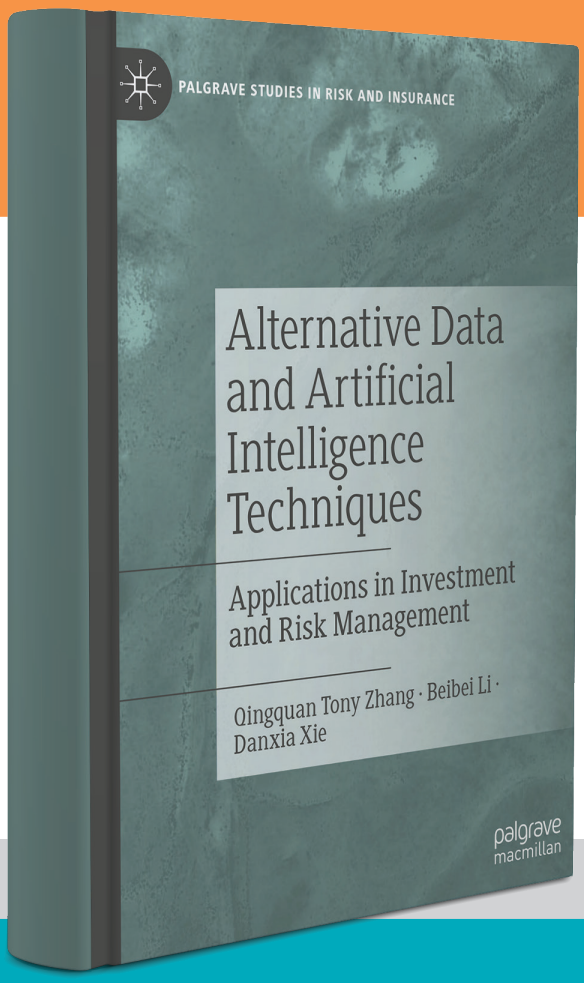


# فراخوان ترجمه کتاب

تمدید شد تا ۲۰ اردیبهشت

پژوهشکده بیمه

**پژوهشکده بیمه، به منظور کمک به گسترش دانش بیمه‌ای، ترجمه کتاب**

Alternative Data and Artificial Intelligence Techniques

Applications in Investment and Risk Management

**را در دستور کار خود قرار داده است. لذا از کلیه اساتید، پژوهشگران، صاحبنظران و کارشناسان دعوت می‌شود که در صورت تمایل به ترجمه کتاب مذکور، کاربرگ درخواست ترجمه پیوست را به همراه سوابق علمی و اجرایی خود و ترجمه صفحات ذکر شده با ذکر عنوان کتاب، حداکثر تا تاریخ ۱۴۰۲/۰۲/۱۵ به آدرس ایمیل nashr@irc.ac.ir ارسال فرمایند.**

| معیارهای ارزیابی | امتیازات | ضریب |
|---|---|---|
| کیفیت ترجمه | میانگین امتیاز ۲ داور (حداکثر ۱۰) | ۱ |
| سوابق علمی | سوابق علمی مرتبط با موضوع کتاب: دکترا ۱۰ – ارشد ۸ – کارشناسی ٦<br>سوابق علمی غیرمرتبط: دکترا ٤ – ارشد ۳ – کارشناسی۲ | ۰.۲ |
| سوابق تالیف/ترجمه کتاب | سوابق مرتبط با موضوع کتاب: حداکثر ۱۰ امتیاز براساس نرمال‌سازی<br>سوابق غیرمرتبط : ۲۰ درصد امتیاز فوق | ۰.٤ |
| سابقه فعالیت تخصصی در حوزه بیمه | حداکثر ۱۰ امتیاز براساس نرمال‌سازی | ۰.٤ |

# کاربرگ درخواست ترجمه کتاب

**عنوان کتاب:**

Alternative Data and Artificial Intelligence Techniques
Applications in Investment and Risk Management

## الف- اطلاعات عمومی

| | |
|---|---|
| نام و نام خانوادگی | |
| شغل و سمت فعلی | |
| مرتبه علمی (ویژه اعضای هیات‌علمی) | |
| آخرین مدرک تحصیلی و رشته | |
| آدرس | |
| شماره تماس ثابت | |
| شماره تماس همراه | |
| پست الکترونیک | |

## ب- سابقه تالیف/ترجمه (حداقل ۳ عنوان از آثار خود را اعلام بفرمائید)

| ردیف | عنوان کتاب/ترجمه | سال انتشار | ناشر |
|---|---|---|---|
| | | | |

## ج- سابقه اجرایی

| ردیف | محل خدمت | مدت زمان خدمت |
|---|---|---|
| | | |

such as neural networks are also utilized to predict portfolio returns and build investment models (Freitas et al. 2009).

## 1.2 Types of Portfolio Management

Even though the asset management industry is changing rapidly due to advances in technology and the expansion of financial markets, asset management is typically categorized as either "passive" or "active", based on asset management managers' styles.

Passive Management Portfolios

A passive management portfolio refers to holding a diversified portfolio with pre-defined holding criteria (e.g., the S&P 500 tracks the largest 500 companies based on market-cap). No other resources or research efforts are needed to analyze securities to improve investment performance once the criteria are specified and publicized. According to the market efficiency hypothesis, if markets are efficient and prices reflect all relevant information, it could be beneficial to adopt a passive strategy without wasting resources trying to figure out competitors in financial markets.

Normally, the managers of passive portfolios consider it practically impossible to outperform the market, and therefore, they limit themselves to tracking it.

Passive management can be divided into two categories: one is the strategy of building a bond portfolio to obtain sufficient funds to repay a debt in future. This is called "immune strategy under single-payer debt", also known as "interest rate disinfection"; the other is the bond portfolio strategy established in order to secure sufficient funds to repay each debt in the future debt flow, which is called either "immune strategy under multiple payment liabilities" or "cash flow matching strategy".

Active Management Portfolios

Active management refers to when bond investors strive to choose the right market opportunity to adjust their portfolio through the prediction and analysis of the general trends of market interest rate changes, so as to minimize risk and maximize returns.

The central idea is that when interest rates change, there will inevitably be incorrectly priced bonds in the market. Prior analysis

has shown that buying these bonds can mitigate the risk of interest rate changes.

These portfolio managers believe that it is possible to earn a higher return than the market, and therefore, with more active portfolio management, in theory, they are able to obtain Alpha (the excess return with respect to the market). For comparative purposes, it is also important to know the Beta of the portfolio, as this indicates its deviation with respect to the market.

Faced with such a variety of styles and assets, portfolio management techniques allow us to standardize management techniques and determine adequate measures of performance and risk, and therefore, create portfolios to suit each type of investor.

Since any portfolio consists of assets, let's define some of the most common assets in portfolios.

## 1.3 The Classic Asset and Derivatives

### *1.3.1 Classic Assets Classes in Portfolio Management*

The classic components of a portfolio are:

Stocks or Equities: These assets are considered high risk and therefore offer opportunities for high returns. They can also be classified by country, sector, and/or as value or growth; the former tends to offer dividends, while the latter tend not to offer dividends, but the return is usually higher than the market. In addition, they offer rights at shareholders meetings, since in fact, the shareholders own the company. (Sometimes there are exceptions with class A and class B shares where the latter do not offer voting rights)

Bonds: These assets are considered low-risk and therefore offer moderate returns; the advantage is that from the outset, the investment return is known. These assets are loans that are made to governments and companies; the maturity date and fixed return or coupon payment at a certain frequency is fixed. They can also be classified as either corporate or government bonds, or by maturity date, rating, etc.

Cash: We must also hold some money in our portfolio because it will not always be 100% invested. Sometimes the cash can be used for operations in the money market, and the risk can vary from very low to very high, depending on the chosen currency.

### *1.3.2 Derivatives in Portfolio Management*

In derivative products, as their name implies, the value (and therefore the return) is derived from an underlying product. Some of the best-known derivatives are:

#### Futures

A future is a contract between two parties, where one party buys/sells an asset with a certain future delivery date and the other party sells/buys the asset. This contract is guaranteed by a clearinghouse.

A future's price is fixed at the transaction time. Thus, at maturity, one party is obliged to deliver the good regardless of the current market price, and the other is obliged to receive the good regardless of the market price, at the current price. The future price is derived from an underlying asset such as a stock index or a commodity.

The most common assets underlying futures include:

- Indices (Dow Jones, S&P 500, Nasdaq, Russell 2000, etc.)
- Precious metals (gold, platinum, etc.)
- Industrial metals (copper, lead, etc.)
- Energy (oil, natural gas, ethanol, etc.)
- Agriculture (soybeans, beans, corn, wheat, etc.)
- Soft commodities (meat, live cattle, orange juice, sugar, cotton, etc.).

#### Options

Options are instruments where one party buys the right to buy/sell an asset at a fixed price on a certain date and the other party sells that right committing to it. This operation is guaranteed by a clearinghouse. By definition, the options give the buyer the right, but not the obligation, to buy or sell the asset. The option price is derived from an underlying asset such as stocks, stock indexes, or commodities. Unlike equities, an option does not represent ownership in the underlying company, though it does give investors ownership once exercised.

There are two categories of options: call options and put options. Call options give the buyer the right to buy the specified asset/equity at the

CHAPTER 10

# Fraud and Deception Detection: Text-Based Data Analytics

With the trend of increasingly complex big data, how to handle and improve the authenticity of data has become an important issue related to the credibility of data. This chapter discusses how to imitate and detect similar applications and how to identify fake reviews by machine learning and various statistical methods using deceptive applications and fake reviews as examples.

Text mining in big data analysis is emerging as a powerful tool to uncover fraud or deception. Since it is of critical importance to many including law enforcement and security personnel (Fuller, Biros, and Delen 2011), by analyzing unstructured text data, text-based data analytics can extract new knowledge, identify significant patterns, and find correlations hidden in the data (Hassani et al. 2020). It is a very effective method of detecting fraud and deception.

Text-based data analytics use text mining algorithms. There are two types of them: supervised learning and unsupervised learning. Supervised learning algorithms use the target's observed values to build a prediction model while unsupervised learning algorithms use a set of predictors to reveal hidden structures in the data (Guduru 2006). Opinion classification and sentiment classification are two approaches extensively used in Text-based data analytics (Esuli 2006). We will see them in the following examples.



Q. T. Zhang et al., *Alternative Data and Artificial Intelligence Techniques*, Palgrave Studies in Risk and Insurance,
https://doi.org/10.1007/978-3-031-11612-4_10

## 10.1 Copycat Detection

Quan Wang, Beibei Li, and Param Vir Singh from CMU used machine learning techniques such as natural language processing, latent semantic analysis, network-based clustering, and image analysis, to analyze whether and how copycats affect an original app's demand. We will then follow their approach to detail the specific application of text-based data analysis on Copycat Detection.

First, we define copycat apps as those that provide similar functions to the original app but that are released at a later date. Then, it further classifies two types of copycat apps—deceptive and nondeceptive. A deceptive copycat tries to deceive the customer into thinking it is the original app by choosing an app name and app icon very similar to that of the original. By contrast, a nondeceptive copycat tries to differentiate itself from the original by choosing a name and icon very different from that of the original.

Through a new copycat detection method that combines a variety of statistical and machine learning methods, it can empirically identify copycats on a large scale based on both functionality and appearance. It can also tell if an app is deceptive. The dataset used in this study is publicly available app information from the US iOS store for the iPhone. It consists of a sample of 10,100 action games by 5,141 developers released between July 2008 and December 2013. The data used in detection includes the app description, user reviews, release date, developer's Apple ID, app name, app icon image, and much more.

The specific detection process is as follows. The first step is to detect functional similarities between apps based on textual descriptions and customer reviews. First of all, it converts the textual descriptions and consumer reviews of all of the apps into a bag of words. Then it performs text preprocessing, such as part-of-speech tagging. Second, the detection process keeps the nouns and verbs because nouns and verbs are more relevant to the app's function. Finally, a TF-IDF matrix is used to identify the functions of the apps to calculate the functional similarity between applications combined with SVD. TF-IDF is a statistical method for creating a term-document matrix. The main idea is that if a word appears frequently in the text converted from an app but appears rarely in other apps, the particular word count is inflated for classification. In this paper, the word can be used to represent the app's function. After identifying the keyword