Saeede Sadat Asadi Kakhki

Progress Report for fall 2016 Semester

PROGRESS

In the following I will briefly list what I have done so far:

1. Explored TMX dataset (trades and quotes fields)
2. Familiarized myself with stock marketing vocabulary
3. K-NN classification and Naïve Bayes algorithm: Explored various datasets such as casandie, sdrapcty, TitanicSurvival, HouseVotes84, iris
4. Read "Portfolio Selection" paper suggested by Hamilton [1]
5. Read "Procedures for Performing Systematic Reviews" paper [2]
6. Familiarized myself with Bloomberg [3]
7. Took "Data Analytics: Basic Methods", "Database Organization for Data Analysts", and "Machine Learning" courses
8. Graduate Assistant for "Information systems" course
   Helped me to learn and review some tools such as MS Excel and Access and their features
9. Graduate Assistant for "Introduction to Big Data"
   Helped me to learn different tools such as SAS, IBM Watson, and Weka
10. Learned working with tools such as R, MySQL Workbench
11. Learned an online course from Pluralsight website: Algorithms and Data Structures 1
12. Learned an online course from Pluralsight website: Algorithms and Data Structures 2
13. Learned an online course from Pluralsight website: Learning to Program (Python)

WORK PLAN

In the following I will list what I am going to do in the rest of this semester:

1. Improving my Python skill by learning "Understanding Machine Learning with Python" online course from Pluralsight website
2. Searching and implementing appropriate Machine Learning algorithms on top of Bloomberg and TMX data
3. Studying more about the works already done in TMX project and watching its presentations
4. Explore Bloomberg and TMX data by appropriate R approaches
5. Trying to study a number of the most relevant Papers from Github-TMX group
6. Visiting Bloomberg University (By typing <BU> into the Bloomberg system, the user will be taken to a screen that offers a wide variety of training videos, informational papers and live seminars)
7. Reading "A Systematic Literature Review on the Applications of Bayesian Networks to Predict Software Quality" paper
8. Reading "Emerging Technologies for Improving Embryo Selection: A Systematic Review" paper
9. Cooperating with Parisa on Systematic Review plan

APPENDIX

[1] Portfolio Selection

Author: Harry Markowitz

THE PROCESS OF SELECTING a portfolio may be divided into two stages. The first stage starts with observation and experience and ends with beliefs about the future performances of available securities. The second stage starts with the relevant beliefs about future performances and ends with the choice of portfolio. This paper is concerned with the second stage. We first consider the rule that the investor does (or should) maximize discounted expected, or anticipated, returns. This rule is rejected both as a hypothesis to explain, and as a maximum to guide investment behavior. We next consider the rule that the investor does (or should) consider expected return a desirable thing and variance of return an undesirable thing. This rule has many sound points, both as a maxim for, and hypothesis about, investment behavior. We illustrate geometrically relations between beliefs and choice of portfolio according to the "expected returns-variance of returns" rule.

[2] Procedures for Performing Systematic Reviews

The objective of this report is to propose a guideline for systematic reviews appropriate for software engineering researchers, including PhD students. A systematic review is a means of evaluating and interpreting all available research relevant to a particular research question, topic area, or phenomenon of interest. Systematic reviews aim to present a fair evaluation of a research topic by using a trustworthy, rigorous, and auditable methodology. The guideline covers three phases of a systematic review:

- Planning the review
    - Identification of the need for a review
    - Development of a review protocol.
- Conducting the review
    - Identification of research
    - Selection of primary studies
    - Study quality assessment
    - Data extraction & monitoring
    - Data synthesis
- Reporting the review

It is at a relatively high level. It does not consider the impact of question type on the review procedures, nor does it specify in detail mechanisms needed to undertake meta-analysis. All these guidelines are intended to aid medical researchers. This document attempts to adapt the medical guidelines to the needs of software engineering researchers. Individual studies contributing to a systematic review are called *primary* studies; a systematic review is a form a *secondary* study.

Reasons for Performing Systematic Reviews

There are many reasons for undertaking a systematic review. The most common reasons are:

- To summarise the existing evidence concerning a treatment or technology
- To identify any gaps in current research in order to suggest areas for further investigation
- To provide a framework/background in order to appropriately position new research activities.

In particular, researchers performing a systematic review must make every effort to identify and report research that does not support their preferred research hypothesis as well as identifying and reporting research that supports it.

Advantages and disadvantages

Systematic reviews require considerably more effort than traditional reviews. Their major advantage is that they provide information about the effects of some phenomenon across a wide range of settings and empirical methods. If studies give consistent results, systematic reviews provide evidence that the phenomenon is robust and transferable. If the studies give inconsistent results, sources of variation can be studied. A second advantage, in the case of quantitative studies, is that it is possible to combine data using meta-analytic techniques. This increases the likelihood of detecting real effects that individual smaller studies are unable to detect. However, increased power can also be a disadvantage, since it is possible to detect small biases as well as true effects.

Feature of Systematic Reviews

Some of the features that differentiate a systematic review from a conventional literature review are:

- Systematic reviews start by defining a review protocol that specifies the research question being addressed and the methods that will be used to perform the review.
- Systematic reviews are based on a defined search strategy that aims to detect as much of the relevant literature as possible.
- Systematic reviews document their search strategy so that readers can access its rigour and completeness.
- Systematic reviews require explicit inclusion and exclusion criteria to assess each potential primary study.
- Systematic reviews specify the information to be obtained from each primary study including quality criteria by which to evaluate each primary study.
- A systematic review is a prerequisite for quantitative meta-analysis.

Finally, these guidelines are intended to assist PhD students as well as larger research groups. However, many of the steps in a systematic review assume that it will be undertaken by a large group of researchers. In the case of a single research (such as PhD student), we suggest the most important steps to undertake are:

- Developing a protocol
- Defining the research question
- Specifying what will be done to address the problem of a single researcher
- Applying inclusion/exclusion criteria and undertaking all the data extraction

- Defining the search strategy
- Defining the data to be extracted from each primary study including quality data
- Maintaining lists of included and excluded studies
- Using the data synthesis guidelines
- Using the reporting guidelines

[3] Bloomberg

Bloomberg delivers business and markets news, data, analysis, and video to the world, featuring stories from Businessweek and Bloomberg News.

Bloomberg L.P. provides financial software tools such as an analytics and equity trading platform, data services, and news to financial companies and organizations through the Bloomberg Terminal (via its Bloomberg Professional Service), its core revenue-generating product.

Below are just some of the features a Bloomberg Financial Markets Lab offers:

- Cross-asset, real-time data
- Historical time series analysis in all markets including Equity, Fixed Income and FX
- Powerful API integration with Excel®
- Public and private company accounting data and corporate finance dataset
- Portfolio and risk analytics and historical back-testing
- Peer-ranked, customizable investment simulations
- Global economic data for macroeconomic analysis
- Global M&A and IPO database

Bloomberg connects rather seamlessly with Excel, so you can use spreadsheets to analyze data downloaded from Bloomberg. Even better, you can build a spreadsheet that automatically updates your data each time you open it, saving you the trouble of doing so by hand. Best of all, Bloomberg offers sample spreadsheets for common types of analysis (i.e. tracking stock price movements or analyzing company balance sheets), thereby saving you the trouble of building your own spreadsheet. You can also further customize these generic spreadsheets that Bloomberg offers in order to fit your specific needs. If you are interested in incorporating Excel spreadsheets into Bloomberg, type <DAPI> to access the Excel main menu on Bloomberg.