Motivation letter

My name is Alexey Grigorev, and I apply to the **EIT ICT Labs Master's Program in Distributed Systems and Services (DSS)** for the fundamental study of distributed systems, service technologies and cloud programming models. Not only is the technical major attractive for me, but the Innovation and Entrepreneurship Minor is also appealing. This program will definitely help me reveal my potential and materialize my entrepreneurial ideas. In this letter I would like to propose one of my most significant projects: a system for predicting prices based on people's opinion collected throughout the web.

Summary

"Stock exchange mood analyzer" is an artificial intelligence application for predicting probable price directions on stock exchange markets by doing extensive analysis of the content of financial newspapers, magazines as well as blog posts, Twitter and Facebook activity. The predictions are afterwards compared against real data, thus leading to a constant cycle of learning on mistakes and improving accuracy of hypotheses.

The product is a web portal which can monitor the changes of selected stock ticker symbols and inform the predicted direction for each. The ultimate goal is to provide a real-time, reliable, fault-tolerant and scalable service with access from everywhere.

The platform is targeted mainly to financial institutions and investment banks, but it also can be used by individual traders and investors.

Such systems are apparently in an increasingly high demand. The reason for this is that traditional statistical analysis and mathematical modeling methods have failed to predict stock prices under crisis circumstances, when essentially it is mood of people that makes prices go up or down, not the previous history of trades.

The idea has immense potential for further development. One option is to extend the system on the level of countries by analyzing rates on the foreign currency market and suggesting the price directions based on what a country is doing. For example, an armed conflict may mean a decrease in the rates of this country's currency.

Details

Many traders and investors when playing on the marked rely on newspapers and financial magazines, sometimes even on significant people in Twitter, Facebook or Google Plus. Based on opinions from these sources, they decide whether to invest in a certain share or not. Therefore, in general, "bad mood" about a company is likely to lead to a price decline as more people will be interested to sell their shares, and vise versa, if a company is praised, "good mood" will cause increases as more and more people are interested to invest in the company.

However, this activity can be automated, and a machine is able to be taught to analyze the mood found in these sources and to discover whether the content is positive or negative. Constantly learning and testing hypotheses against real data from stock exchanges, eventually it will be possible to find strong patters between the mood and real trading data. The more data processed, the more accurate predictions are made, and over time it is expected to react appropriately in real time environment.

For example, when recently UBS announced job cuts[1], it attracted many investors[2] [3], and, as the result, the price has been growing since October 2012[4]. With software like this all information sources can be processed automatically, with no human interaction at all.

Utilization of Twitter and Facebook plays essential role in price prediction since it reflects the mood of people better than any newspaper. For example, using Twitter it was possible to predict Facebook's IPO results quite accurately[5]. Additionally, there are on-going researches into using social media for predicting stock exchange prices[6] [7], and their results, unless patented, can be used in the project.

The key technical areas in the project are machine learning and artificial intelligence, but in order to be successful, the system is expected to process prodigious amounts of information quickly and effectively. Keeping track only on Twitter may require a cluster of machines, let alone other information sources. Therefore, a distributed fault-tolerant system for handling big data is needed. Designing such a system requires knowledge about parallel and distributed calculation, cloud programming models and big data.

As the goal is to be easily accessible to the end user, it has to be a SaaS system, working on a cloud and accessible from everywhere day and night.

Market analysis

Stock exchanges prices analyzers and predictors have existed for a long time. There are plenty of them[8], including GMDH shell[9], AIQ Systems Trading Expert Pro[10], to name but a few.

In many cases, the basic idea in this kind of software is statistical analysis of past trading activities. While, in general, this is an old and proven method, it fails to predict prices under certain circumstances, including crisis situations and some strategies of the top managers of companies[11].

Additionally, on top of that, most of them are desktop applications, not web-based cloud services.

Current opportunities are used by business intensively. For example, in UBS, Deutsche Bank or any other investment financial institution traders rely on these programs or inner solutions when trading on stock exchanges. Therefore, the key players are likely to invest in the project.

Vision

First and foremost, 0.1 version has to be released. Depending on the number of people involved and their skills it could take up to 2 years – for one person with not all needed knowledge, skills and experience. At this stage, the project requires extensive research.

---

[1]http://www.reuters.com/article/2012/10/30/us-ubs-restructure-idUSBRE89S0DM20121030
[2]http://goo.gl/8bvtT
[3]http://wbponline.com/Articles/View/9545/ubs-shares-rise-on-10k-job-cuts-in-restructuring
[4]http://goo.gl/RSsCE
[5]http://www.kernelmag.com/comment/opinion/2355/facebook-twitter-correlation/
[6]http://www.mediabistro.com/alltwitter/twitter-predict-stock-market_b19796
[7]http://bus.miami.edu/umbfc/_common/files/papers/Karabulut.pdf
[8]http://dmoz.org/Business/Investing/Research_and_Analysis/Software
[9]http://www.gmdhshell.com/
[10]http://www.aiqsystems.com/
[11]The Signal and the Noise: Why So Many Predictions Fail – but Some Don't, Nate Silver

It is important to acquire needed knowledge, skills, make connections, collect data and eventually write a master dissertation on the obtained results.

The next years should be aimed at calibrating the system to making better predictions, collecting more data, and looking for investors. As the result at this stage, version 1.0 is expected as well as some customers and interested investors. Once investors appear, the system can be tailored to their needs and to the needs of business in general.

As for individuals, a limited subscription with possibility to follow only a few stock symbols can be provided for free, in addition to a premium paid subscription with access to information with no limitations. The key factors on prices will be prices offered by companies with similar services.

Finally, in five years it is expected to be a well-established product with its loyal customers plus key players such as financial institutions interested in the product. At this stage the aim is to provide high quality service and deliver results as soon as information appears.

Study objectives

The EIT ICT Labs Master's Program in Distributed Systems and Services (DSS) is the best place for the inception of this project, for its curriculum meets perfectly its and my needs. Distributed Systems, Service Technologies and Cloud Programming Models disciplines are critical for the project. Additionally, Machine Learning, Web Data & Knowledge, Statistics and Data Integration & Semantic Web subjects are of equal importance, and that is why I would like to choose the **Distributed Information Management** specialization.

What really distinguishes the EIT ICT Labs program from others is Innovation & Entrepreneurship Minor. It is not only crucial for this particular project, but also important for my future career in general, as I am full of ideas and it will give me a perfect momentum to put them into actions. I am interested in Business Development, because I want to drive the market, not to be driven by it.

To summarize, with the motivation to materialize my project, and the desire to drive innovations, I am confident that I am qualified and able to perform well in this program.

Thank you very much for considering my application. I am looking forward to your positive response.