

Machine Learning Algorithms III

Module 6, 2017-2018

Ilya Munerman
Munerman & Partners
ivm@munerman.ru

TAs: TBA

Course description

This course describes how dramatically changes of the information market provides new power tools for financial data processing and analysis. In addition, we will compare classical and contemporary tools performance. The main aim of this course is a species of different country data sets and technics for integration this data for common international data environment.

The course is designed for listeners which known elementary economics, finances, IT and mathematics and may be able for economists, IT specialists, managers, include MBA and journalists.

Course requirements, grading, and attendance policies

Statistics, mathematics, corporate finance, assets valuation.
The course grade is based on the final exam (100%).

Course contents

1. Contemporary financial analyses main challenges.

Ontology and its role in the Contemporary financial analyses. Main categories: value, sustainability, dynamic. Understanding a difference between pattern recognition and prediction technics. Data quality and cognitive biases. Expert's professional crisis caused by data revolution.

2. Big, data mining, data science.

Why general universe processing most powerful, then sample analysis? What is meaning sample today? Data source survey. Big data infrastructure. Low and ethical problem using big data.

Main data types: open/close, government/private/corporate, free/paid, numerical/semantic. Data gathering, storage and retrieval instruments. Cloud technologies. Data base management, API, BI, ERP, final processing systems. Providers survey. From state to start-ups. Block chain technology.

3. Main data types.

Quote data for technical and fundamental stock market analysis. Labor data. Real estate data. Procurement and Contracts electronic trading platforms data. Legal and court information. Web-sales and traffic data. Providers, aggregators, efficient and inefficient solutions. Open data hubs: pro et contra. Different solution costs.

4. Correlation.

Correlation vs causalities. Metrics minding. Linear and nonlinear dependencies.

5. Data processing. Modeling.

Recognition, prediction and other forms of data processing. From regression to a neural network. Data mining. Cluster analyses and dendrograms. Stochastic processes. Probability vs reliability.

6. Estimation and model testing.

Statistical tests and criterions: R2, ROC-curve, type I and type II errors, graphical analyses. Ex-post testing.

7. Applications.

Scorings, rankings, ratings. Ordinal and cardinal utilities. Tools for modern institutes: crowdsourcing, start-ups, franchising and so on.

Credit conveyors. Value of credit and collateral value. Credit value monitoring according to regulation and standards.

Real estate apps and valuation. Ad valor tax rates. Real estate databases. State and corporate real estate data hubs.

Collecting data about all types of procurement: government, commercial, international, planned purchases. Data standards problem. Auction houses data. Open societies, economic forecast, parameters estimation applications and so on.

Visualization, infographics and data journalism. Importance, examples, tools, and applications. What is to be done.

Contests and hackathons for provide original salvation.

Course materials

Required textbooks and materials

1. James G., Witten D., Hastie T., Tibshiriani R. (2015) An introduction to statistical learning with applications in R, 6th edition, Springer
2. Brooks C. (2014) Introductory Econometrics for Finance, Third Edition, Cambridge University Press.

Additional materials

1. D&B World bankruptcy reports
2. Certified fraud examiners exam questions
3. Special course materials will be provided in cloud storage

Academic integrity policy

Cheating, plagiarism, and any other violations of academic ethics at NES are not tolerated.