dr Piotr Wójcik
pwojcik@wne.uw.edu.pl
http://www.wne.uw.edu.pl/pwojcik/

Master's seminar „Machine learning in finance and other applications"

Potential topics of master theses:

1. **Measuring regional and local inequalities from space for Poland or Europe** – using data on night-time lights intensity on Earth (pictures from NASA satellites)
2. **Home Credit Default Risk. Can you predict how capable each applicant is of repaying a loan?** – based on kaggle competition: https://www.kaggle.com/c/home-credit-default-risk, huge dataset, diversified machine learning tools can be applied – random forests, support vector machines, xgboost, model ensembling.
3. **Credit scoring in P2P lending** – based on kaggle data: https://www.kaggle.com/skihikingkevin/online-p2p-lending or https://www.kaggle.com/wendykan/lending-club-loan-data
4. **Predicting customer revenues** based on Google Store data – from kaggle competition: https://www.kaggle.com/c/ga-customer-revenue-prediction, diversified machine learning tools can be applied – random forests, support vector regression, xgboost, LASSO, ridge, model ensembling.
5. **Using news to predict stock movements** – based on kaggle competition: https://www.kaggle.com/c/two-sigma-financial-news, diversified machine learning tools can be applied – random forests, support vector machines, xgboost, neural networks, model ensembling.
6. **Predicting house prices with advanced regression techniques** – based on kaggle competition: https://www.kaggle.com/c/house-prices-advanced-regression-techniques, diversified machine learning tools can be applied – random forests, support vector regression, xgboost, LASSO, ridge, model ensembling.
7. **I just run one LASSO regression. Factors of real economic convergence** – a reply to Sala-i-Martin X. (1997b) I just ran two million regressions, American Economic Review, 87, p. 178-183.
8. **The impact of Federal Open Market Comitee minutes on financial markets** in high-frequency domain, constructing and backtesting an event arbitrage strategy, text mining, sentiment analysis of FOMC minutes and checking its impact on predictability of financial markets
9. **Simple combining forecasts vs models ensembling** (stacking/blending) – what works better, case studies on two different datasets – can be financial, but also any other
10. Application of Least Squares Support Vector Machines (LS-SVMs) in prediction of stock prices movements / automated trading strategy applied to Polish / European / Asian financial markets
11. Application of **deep learning** in prediction of stock prices movements / automated trading strategy applied to Polish / European / Asian financial markets
12. Topic related to **image/voice recognition**
13. Topic related to **blockchain** methodology
14. **ANY OWN TOPIC** in which selected machine learning tools (predictive modelling – classification or regression) are used on real data (e.g. from kaggle competitions).

dr Piotr Wójcik
pwojcik@wne.uw.edu.pl
http://www.wne.uw.edu.pl/pwojcik/

Master theses written under my supervision and successfully defended so far

In English

1. Can Twitter predict stock returns?
2. Index Tracking with Post Constrained LASSO. Applications to Warsaw Stock Exchange indexes
3. Index tracking using genetic algorithms
4. Application of neural networks in algorithmic trading
5. Predicting Chinese stock market with Artificial Neural Networks and hybrid models
6. Automated pairs trading on S&P500 stocks using volatility switching regimes
7. Event Arbitrage Strategies. The impact of the U.S. markets opening on European stock indices
8. Optimization of the portfolio of Polish stocks using the Black-Litterman model
9. Investment strategy optimization with the use of genetic algorithms. The analysis for Chinese stock market
10. Application of breadth indicators in automated trading strategies
11. Microstructure of the foreign exchange market and its effects in intraday volatility
12. Construction of volatility index for China CSI300 stock index by using GARCH-type models
13. Application of artificial neural networks in high frequency trading
14. Seeking for momentum in US industrial indices – application of automated trading strategies
15. Commodity-currency trading. Empirical strategies on high-frequency data
16. Co-integration in pair trading strategies. Application to high-frequency data on American equities
17. Using a Bayesian Network to Boost a Trading Strategy: evidence from EURUSD currency pair
18. Portfolio Management based on Statistical Arbitrage Strategy
19. Comparison of predictive abilities of Extreme Value Theory and AR-GARCH approaches in VaR forecasting. Application to Polish stock market
20. Is technical analysis still alive? Application of selected methods in an automated trading strategy
21. Pairs trading strategies on US equities. Sectoral high-frequency data approach
22. Trend following or mean-reversion? Analysis for – Foreign Exchange Market using high frequency data
23. Analysis of High Frequency strategies based on different types of moving averages
24. Profitability of algorithmic strategies applied on high-frequency data
25. Trading strategy for futures contracts based on Hurst exponent

W języku polskim

1. Wykorzystanie analizy falkowej w automatycznej strategii inwestycyjnej na danych o wysokiej częstotliwości
2. Determinanty współczynnika urodzeń w Polsce w ujęciu lokalnym
3. Wzorzec konwergencji bezrobocia w ujęciu stochastycznym – Polska 2004-2016
4. Co wpływa na frekwencję wyborczą? Analiza z wykorzystaniem metod ekonometrii przestrzennej
5. Pierwsza dekada w Unii Europejskiej – sukces czy porażka nowych państw członkowskich?
6. Reakcja polskich konsumentów na kryzys. Wydatki konsumpcyjne gospodarstw domowych w latach 2007–2009
7. Credit scoring jako metoda oceny ryzyka kredytowego osób fizycznych z wykorzystaniem technik data mining
8. Atrakcyjności lokalizacji a wycena lokalu mieszkalnego na przykładzie warszawskiego wtórnego rynku mieszkaniowego
9. Analiza zróżnicowania poziomu rozwoju społeczno-gospodarczego powiatów i gmin województwa małopolskiego ze szczególnym uwzględnieniem powiatu gorlickiego
10. Dojazdy do pracy a konwergencja regionalna w Polsce
11. Fundusze strukturalne a rozwój lokalny na przykładzie gmin województwa wielkopolskiego i podkarpackiego
12. Poziom rozwoju społeczno-gospodarczego powiatu opoczyńskiego na tle powiatów Polski – analiza porównawcza
13. Wpływ atrybutów mieszkania na jego cenę ofertową na przykładzie ogłoszeń z serwisu dom.gratka.pl
14. Analiza przestrzenna polaryzacji i dyfuzji rozwoju regionalnego Polski
15. Analiza konwergencji regionalnej w Polsce i Hiszpanii
16. Prognozowanie liczby pierwszych obciążeń zwrotnych
17. Analiza zróżnicowania poziomu rozwoju społeczno-gospodarczego gmin województwa mazowieckiego
18. Analiza zróżnicowania poziomu życia w regionach Unii Europejskiej