

I Ogólnopolskie Seminarium Naukowe „Uczenie maszynowe w ekonomii i finansach” - UMEF 2021

Streszczenia referatów

dr hab. inż. Grzegorz Dudek

Randomizowane sieci neuronowe jako modele prognostyczne

Sieci neuronowe należą do najczęściej stosowanych modeli predykcyjnych ze względu na swoją elastyczność i własność uniwersalnego aproksymatora. Jednak uczenie sieci neuronowych klasycznymi metodami opartymi na optymalizacji gradientowej jest mało efektywne z uwagi na skomplikowany charakter funkcji błędu. Algorytmy gradientowe utykają w minimach lokalnych i płaskich obszarach tej funkcji, nie gwarantując optymalnych rozwiązań. Proces optymalizacji gradientowej jest złożony i czasochłonny, szczególnie dla skomplikowanych funkcji docelowych, dużych zbiorów uczących i dużej liczby neuronów. Ponadto w architekturach głębokich pojawia się problem zanikających lub eksplodujących gradientów, który dodatkowo utrudnia uczenie się sieci. W sieciach randomizowanych nie stosuje się optymalizacji gradientowej. Parametry neuronów ukrytych są dobierane losowo i nie podlegają uczeniu. Dobiera się jedynie wagi neuronów warstwy wyjściowej. Problem optymalizacyjny staje się wypukły i może być rozwiązany za pomocą standardowej metody najmniejszych kwadratów, co drastycznie redukuje czas treningu sieci.

W referacie przedstawiono model prognostyczny oparty na randomizowanej sieci neuronowej. Aby wzmocnić własności aproksymacyjne i generalizacyjne sieci randomizowanej, parametry losowe generowano uwzględniając dystrybucję danych i złożoność funkcji docelowej. Reprezentacja danych oparta na wzorcach sekwencji szeregów czasowych umożliwia prognozowanie szeregów czasowych z nieliniowym trendem, wielokrotną sezonowością i zmienną wariancją. Prosta i czytelna architektura, szybka optymalizacja i trening, łatwa implementacja i wysoka dokładność czynią proponowany model atrakcyjnym w rozwiązywaniu złożonych problemów prognostycznych.

mgr Małgorzata Grządzielewska

Selekcja zmiennych z wykorzystaniem modelu regresji wielorakiej i wybranych metod uczenia maszynowego

Prawidłowy dobór zmiennych diagnostycznych pozwala na stworzenie modelu cechującego się większą dokładnością przewidywania. Uwzględnienie nieistotnych cech może powodować nadmierne dopasowanie modelu do danych, przez co traci on zdolność generalizacji. Celem pracy jest porównanie metody regresji wielorakiej z wybranymi modelami uczenia maszynowego (tj. LASSO, sieci elastycznych, CART i Lasów Losowych) do selekcji zmiennych objaśniających. Analizę poprzedzono przedstawieniem podstawowych zagadnień związanych z selekcją cech, w tym sposobów i kryteriów doboru zmiennych. Badanie przeprowadzono na podstawie danych dotyczących poziomu wypalenia zawodowego wśród pracowników socjalnych. Następnie, dokonano wyboru optymalnej metody, poprzez ocenę

jakości dopasowania modeli i ich mocy predykcyjnej poza próbą. Zdolność predykcyjną modeli zbadano za pomocą błędu średniokwadratowego obliczonego metodą sprawdzania krzyżowego (MSE_{CV}).

dr Jan Kubacki

On testing the quality of various econometric models under large dynamics

Because it may be interesting to examine whether the observed economic trends will continue in the coming years. Taking into account the current situation related to the coronavirus pandemic and the existence of many factors limiting economic growth (in particular the GDP per capita growth), the results of the presented approach should be treated with special caution. Due to the high dynamics of the currently observed economic phenomena (see Wójcik, 2014), the assessment of such changes on the basis of yearly data may be insufficient. We examine the monthly trends of available income given by Polish Public Statistics and assess the characteristics of this trends using various econometric models i.e. Vector Autoregression Models (VAR), Kalman filter models (using dlm and KFAS R-package), classical economic growth models (for example using panel models including generalized method of moments), and classical linear models (also with non-linear transformations) and taking into account the sliding spans. During the analysis the behaviour of various diagnostics tests was performed. The R-project and GRETl packages were used. In particular the vars, dlm, KFAS, lm and splines packages were used.

dr hab. prof. UMK Witold Orzeszko

Analiza zależności przyczynowych między stopami zmian indeksów giełdowych a wolumenem obrotów

W literaturze przedmiotu wskazuje się teoretyczne przesłanki uzasadniające istnienie dwukierunkowej przyczynowości między stopami zwrotu instrumentów finansowych a wolumenami ich obrotów. Jednak wyniki badań empirycznych są zróżnicowane i nie potwierdzają jednoznacznie jej obecności. W przeprowadzonym badaniu dokonano próby detekcji nieliniowej przyczynowości w sensie Grangera między stopami zmian wybranych indeksów z Giełdy Papierów Wartościowych w Warszawie a wolumenami ich obrotów. Z definicji, obecność przyczynowości w sensie Grangera oznacza możliwość efektywnego prognozowania badanych szeregów czasowych. Jednak z praktycznego punktu widzenia, wiedza o potencjalnej prognozowalności zwykle nie jest celem samym w sobie, lecz punktem wyjścia do poszukiwania skutecznej metody prognozowania. Tymczasem testy nieliniowej przyczynowości mają zazwyczaj charakter nieparametryczny, co oznacza, że weryfikuje się w nich hipotezę zerową o braku przyczynowości względem hipotezy alternatywnej głoszącej, że analizowane procesy powiązane są pewną relacją przyczynową o niesprecyzowanej postaci. W efekcie odrzucenie hipotezy zerowej nie wnosi informacji o postaci analitycznej modelu opisującego zależności między badanymi procesami. Z tego względu naturalne jest aby do prognozowania zidentyfikowanych relacji wykorzystać metody, które nie wymagają założeń o ich postaci analitycznej, lecz pozwalają danym mówić same za siebie, a ponadto są skuteczne do prognozowania nieliniowości. Tego typu własnościami cechują się metody uczenia

maszynowego. W pracy dokonano próby wykorzystania jednej z nich – metody wektorów nośnych jako narzędzia prognozowania analizowanych procesów.

dr Paweł Sakowski, mgr Jakub Michańków, dr hab. Robert Ślepaczuk

The comparison of LSTM in algorithmic investment strategies on BTCUSD and S&P500 index on various frequencies

The use of the LSTM networks to forecast the value of the BTCUSD and S&P 500 index on the data from 2013 to the end of 2020 on the following frequencies: daily, 1h and 15min data. Based on the forecasts from ML models we generate buy and sell investment signals, employ them in algorithmic investment strategies and create equity lines for our investment. We use for this purpose various combinations of ML models optimized on in-sample periods. We pay special attention to data preprocessing in the input layer, to avoiding overfitting in the estimation and optimization process, and correct selection of hyperparameters at the beginning of our tests. The next stage is devoted to the conjunction of signals from various frequencies into one ensemble model and the selection of the best combinations for out-of-sample period through the optimization of the given criterion within the similar way as in the portfolio analysis. Finally, the sensitivity analysis of the main parameters and hyperparameters were performed.

dr inż. Michał Stasiak

Budowa systemu handlu algorytmicznego na podstawie modelowania stanowego z uwzględnieniem analizy trendu

Analiza trendu jest jedną z głównych metod stosowanych przez inwestorów do predykcji kierunku przyszłych zmian instrumentów finansowych. Brak jednak jasno zdefiniowanych zasad detekcji rodzaju, początku oraz końca trend. Większość metod analizy trendu zdefiniowanych w analizie technicznej zakłada zastosowanie popularnej w praktyce inwestycyjnej oraz badaniach naukowych reprezentacji świecowej. Reprezentacja ta prowadzi do utraty informacji o zmianach charakteryzujących się niską amplitudą oraz wysoką częstotliwością (zmiany te są szczególnie ważne w kontekście projektowania systemów HFT). Dokładniejsze rezultaty można otrzymać na podstawie reprezentacji binarno-czasowej. W referacie przedstawiono propozycje budowy modelu stanowego, pozwalającego na precyzyjną detekcję trendu oraz jego parametrów w reprezentacji binarno-czasowej, a także predykcję kierunku przyszłych zmian na podstawie tych informacji. Na podstawie proponowanego modelu zbudowano system HFT oraz przetestowano jego działanie z wykorzystaniem dedykowanego oprogramowania napisanego w językach C++ oraz Mql4. Badania empiryczne przeprowadzono z wykorzystaniem danych historycznych dla rynku złota (notowania instrumentu XAU/USD w okresie 2014-2020).

mgr Marcin Stawarz

Prognozowanie stóp zwrotu indeksów giełdowych z wykorzystaniem modeli uczenia maszynowego

Prognozowanie stóp zwrotu na rynkach finansowych jest bardzo pożądanym i zarazem jednym z trudniejszych zadań. Największym problemem jest mnogość i złożoność czynników wpływających na przyszłą stopę zwrotu instrumentu finansowego. Głównym celem pracy jest sprawdzenie czy modele uczenia maszynowego mogą charakteryzować się większą dokładnością prognoz w porównaniu z klasycznym modelem jakim jest, np. ARIMA. Wykorzystanie metod uczenia maszynowego może pozwolić na znalezienie i wykorzystanie zależności, które mogłyby zostać pominięte przez klasyczne modele ekonometryczne. W tym celu dla wybranych indeksów giełdowych zostały oszacowane trzy modele: ARMA, SVR oraz las losowy. W przypadku każdego z tych modeli przeprowadzono dobór optymalnych parametrów. Następnie bazując na modelach z najlepszymi parametrami zostały obliczone prognozy dla każdego badanego indeksu. Kryterium porównawczym dla wszystkich modeli była dokładność otrzymanych za ich pomocą prognoz, w tym celu zostały wykorzystane następujące miary dokładności prognoz: RMSE oraz MAE.

mgr Aleksandra Szymura

Wykorzystanie metod uczenia nadzorowanego w procesie oceny działalności spółek kapitałowych

Zapobieganie przestępczości gospodarczej i wszelkiego rodzaju nadużyć jest obecnie ogromnym wyzwaniem w gospodarce. Wiele firm w swojej działalności popełnia błędy (świadomie lub też nie), które skutkują koniecznością zapłaty kar, czy też odszkodowań na rzecz innych podmiotów. Może to w ostateczności zachwiać ich płynnością finansową, co w konsekwencji może spowodować nawet ogłoszenie upadłości lub likwidację.

W swojej pracy badawczej autorka porusza problematykę oceny działalności podmiotów z wykorzystaniem metod uczenia nadzorowanego, głównie przy użyciu odpowiednich algorytmów klasyfikacyjnych. Celem jest utworzenie takiego narzędzia, za pomocą którego można będzie ocenić prawdopodobieństwo zapłacenia przez podmiot kary w kolejnym roku. Dzięki powstałemu algorytmowi możliwe będzie podjęcie trafnej decyzji przez instytucje finansowe, np. banki, na temat udzielenia ewentualnego kredytu. Ponadto sprawdzona zostanie ważność poszczególnych cech oraz wybór tych, które faktycznie mogą sygnalizować wystąpienie w przyszłości poruszonego w pracy negatywnego zjawiska.

W pracy pod uwagę wzięto dane pochodzące ze sprawozdań finansowych spółek Rynku Głównego oraz NewConnect Giełdy Papierów Wartościowych w Warszawie. Wykorzystane zostały wybrane metody uczenia nadzorowanego, a także zastosowano klasyfikator głosujący, tzw. *voting classifier*, uwzględniający kombinację wyników wybranych algorytmów klasyfikacyjnych. Efektem końcowym jest wybór metody dającej najlepsze wyniki. Warto jednak zaznaczyć, że temat oceny działalności podmiotów nie został zakończony. W przyszłych badaniach warto rozważyć rozszerzenie warsztatu metod uczenia maszynowego lub przetestować inne zmienne, w tym różne warianty cechy zależnej.

dr hab. inż Krzysztof S. Targiel

Prognozowanie zmienności z wykorzystaniem złożonych modeli neuronowych

Zmienność rynku jest jednym z kluczowych problemów zarządzania ryzykiem finansowym. Wśród różnych metod stosowanych do jego określenia wykorzystywano także uczone maszynowo sieci neuronowe. Uzyskiwane rezultaty nie są jeszcze w pełni satysfakcjonujące. Rodzi to potrzebę poszukiwanie nowych modeli, które zapewnią wystarczający poziom dokładności oszacowania zmienności.

W referacie zostanie przedstawione zastosowanie modeli złożonych (ang. *ensemble models*) do prognozowania zmienności. Bazując na prostych modelach wykorzystujących do prognozowania zmienności ceny otwarcia, zamknięcia oraz różne predyktory zmienności bazujące na tych wartościach, zostaną rozszerzone te sieci neuronowe o inne modele wykorzystujące także inne czynniki mające wpływ na zmienność rynku. Weryfikacja modeli zostanie przeprowadzona na danych dotyczących indeksu VIX.