

dr hab. Agata Kliber, prof. UEP
Katedra Matematyki Stosowanej
Uniwersytet Ekonomiczny w Poznaniu
Al. Niepodległości 10
61-875 Poznań

Poznań, 09.09.2024

Recenzja rozprawy doktorskiej pana magistra Piotra Bednarskiego:
*Wykorzystanie algorytmów NLP do automatycznej analizy treści informacji
finansowych w procesach prognozowania notowań giełdowych*
(dziedzina: nauki społeczne, dyscyplina: ekonomia i finanse)
napisanej na Uniwersytecie Ekonomicznym w Krakowie
pod kierunkiem prof. UEK dr. hab. inż. Janusza Morajdy
oraz dr. Romana Huptasa

Rozprawa pana magistra Piotra Bednarskiego koncentruje się na badaniu możliwości poprawy prognoz cen i zwrotów instrumentów finansowych poprzez uwzględnienie w modelu prognostycznym informacji zawartych w wiadomościach oraz postach w mediach społecznościowych dotyczących tych instrumentów, przy czym informacje te ekstrahowane są z wykorzystaniem metod przetwarzania języka naturalnego.

Teza, cel pracy i pytania badawcze

Celem rozprawy było *zbadanie możliwości zastosowania najnowszych metod przetwarzania języka naturalnego do zautomatyzowanej analizy wiadomości oraz postów o tematyce finansowej, gospodarczej i politycznej, a także uwzględnienia ich w finansowych modelach prognostycznych.*

Cele szczegółowe sformułowane przez Autora, które miały doprowadzić go do realizacji celu głównego, to:

1. Określenie źródeł danych wywołujących najsilniejsze reakcje na rynkach finansowych.
2. Zbadanie możliwości efektywnego wykorzystania głębokich sieci neuronowych oraz innych metod przetwarzania języka naturalnego do automatycznej analizy wiadomości i postów w mediach społecznościowych pod kątem wykorzystania ich do przewidywania dynamiki notowań giełdowych.
3. Zaproponowanie i implementacja autorskich modeli hybrydowych łączących modele klasyczne z automatyczną analizą treści oraz ocena ich efektywności prognostycznej.

4. Zbadanie, w jaki sposób ceny różnych aktywów reagują na informacje na ich temat oraz określenie czasu trwania ewentualnego zaburzenia.
5. Porównanie autorskich metod z rozwiązaniami stosowanymi dotychczas w literaturze.
6. Stworzenie autorskich programów komputerowych w ogólnie dostępnych środowiskach programistycznych.

Hipoteza główna rozprawy nie została przez Autora sformułowana. Doktorant sformułował natomiast siedem hipotez cząstkowych:

1. Możliwe jest efektywne zautomatyzowanie analizy finansowych wiadomości oraz postów w mediach społecznościowych poprzez użycie współczesnych metod przetwarzania języka naturalnego – głównie głębokich sieci neuronowych.
2. Wykorzystanie treści całej wiadomości w analizie pozwala na uzyskanie lepszych (dokładniejszych) prognoz punktowych cen i stóp zwrotu wybranych aktywów finansowych niż w przypadku wykorzystania wyłącznie treści ich nagłówków.
3. Modele oparte na głębokich sieciach neuronowych z autorskim modulem FENE generują lepsze (dokładniejsze) punktowe prognozy cen i stóp zwrotu aktywów finansowych w porównaniu z modelami wykorzystującymi klasyczne metody sztucznej inteligencji.
4. Wykorzystanie szeregu czasowego notowań o 5-minutowej częstotliwości jest wystarczające, aby zaobserwować zmiany wielkości stóp zwrotu analizowanych aktywów finansowych spowodowane publikacją wiadomości.
5. Zautomatyzowana analiza finansowych wiadomości i postów w mediach społecznościowych za pomocą algorytmów NLP w połączeniu z obecnie stosowanymi metodami prognozowania szeregów czasowych pozwala na poprawę efektywności predykcji.
6. Klasyczne metody ekonometrii finansowej typu ARMA-GARCH z autorskim modulem FENE generują lepsze (dokładniejsze) punktowe prognozy cen i stóp zwrotu aktywów finansowych w porównaniu z klasycznymi modelami typu ARMA-GARCH bez tego modułu.
7. Klasyczne metody ekonometrii finansowej typu ARMA-GARCH z autorskim modulem FENE generują lepsze (dokładniejsze) punktowe prognozy cen i stóp zwrotu aktywów finansowych w porównaniu z modelami opartymi na głębokich sieciach neuronowych z autorskim modulem FENE.

Struktura pracy

Pracę rozpoczyna obszerny **Wstęp**, w którym Autor przedstawia cel i hipotezy swojej rozprawy, syntetycznie opisuje wyniki uzyskane w kolejnych częściach pracy oraz podsumowuje najważniejsze wnioski.

W rozdziale pierwszym pt. *Informacja na rynkach finansowych oraz jej wykorzystanie w procesach prognozowania* została zaprezentowana klasyfikacja rynków finansowych, opisane źródła informacji dostępnej dla inwestorów oraz dotychczasowe możliwości jej wykorzystania w modelach prognostycznych.

Rozdział drugi pt. *Modelowanie i prognozowanie jednowymiarowych finansowych szeregów czasowych* to syntetyczne przedstawienie modeli ARMA-GARCH oraz modeli uczenia maszynowego. W **rozdziale trzecim** pt. *Współczesne techniki NLP oraz możliwości ich zastosowania do analizy informacji finansowych* Autor przedstawił historię, rozwój i możliwości zastosowania technik przetwarzania języka naturalnego, ze szczególnym uwzględnieniem zastosowań w kontekście prognoz instrumentów finansowych. Pod koniec tego rozdziału Autor sformułował lukę badawczą, jako *brak podejść wykorzystujących w prognozowaniu wiadomości w częstotliwości śróddziennej*.

Rozdział czwarty pt. *Proponowana metodologia badań wykorzystania technik NLP w analizie treści wiadomości finansowych* poświęcony jest dokładnemu opisowi metod zastosowanych w pracy. Autor zaczyna od powtórzenia hipotez badawczych przedstawionych we wstępie rozprawy i jednocześnie wskazuje sposoby ich weryfikacji. Uważam to za bardzo trafne podejście, ponieważ pozwala zminimalizować niejasności związane z postawionymi wcześniej hipotezami (o których piszę w części *Kwestie dyskusyjne i pytania do Autora*). Doktorant opisuje wyczerpująco architekturę własnego systemu prognozowania notowań giełdowych, zaczynając od sposobu przygotowania danych, aż do estymacji i oceny modeli. W podrozdziale 4.5 opisuje własności statystyczne wykorzystanych szeregów czasowych i uzasadnia ich wybór.

W **rozdziale piątym** *Badanie efektywności predykcji finansowych szeregów czasowych wspomaganej narzędziami NLP* Doktorant przedstawia badanie empiryczne, w którym wykazuje wyższość autorskiego podejścia do prognozowania szeregów czasowych ponad modele ARMA-GARCH. Na podkreślenie zasługuje szczegółowy i dokładny opis przeprowadzonego badania (w tym wyboru najlepszego modelu) oraz rysunki 5.2-5.11 przedstawiające, jak informacja odzwierciedlana jest w zwrotach w miarę upływu czasu, poprzez porównanie błędów prognoz dla modeli z informacją i bez. Bardzo trafnym pomysłem jest również zweryfikowanie, która informacja rzeczywiście wywiera wpływ na zmiany stóp zwrotu, a która jest szumem (zróznicowanie źródeł informacji – podrozdział 5.3).

Rozdział szósty pt. *Architektura modeli hybrydowych ARMA-GARCH-FENE oraz analiza ich efektywności* poświęcony jest badaniu efektywności prognoz modeli ARMA-GARCH, do których jako dane egzogeniczne dostarczany jest szereg informacji pochodzących z danych tekstowych, przetworzony przez moduł FENE. Autor porównuje w nim jakość prognoz takiego modelu z jakością prognoz modelu ARMA-GARCH oraz LSTM-FENE, wykazując między innymi, że model hybrydowy ARMA-GARCH-FENE dawał lepsze prognozy (w sensie statystycznym) tylko w przypadku czterech akcji, natomiast w przypadku pozostałych instrumentów lepszy lub równie dobry okazał się model ARMA-GARCH.

Rozdział siódmy stanowi podsumowanie wyników uzyskanych w rozprawie. Oceniam go bardzo wysoko. Stanowi on trafne i syntetyczne podsumowanie procesu badawczego. Autor przypomina hipotezy badawcze oraz wyniki badań przeprowadzonych w celu ich weryfikacji. Rzeczowo przedstawia ograniczenia swojego podejścia i omawia możliwe kierunki jego rozbudowy. Dodatkowo do pracy jest **Aneks** zawierający listing kodu oraz dodatkowe obliczenia, których Autor ze względu na przejrzystość tekstu nie umieścił w głównej części pracy.

Najważniejsze osiągnięcia Autora

Najważniejszymi osiągnięciami Autora są:

- Zaproponowanie i implementacja autorskich modeli hybrydowych łączących modele klasyczne z automatyczną analizą treści.
- Zbadanie, w jaki sposób ceny aktywów różnych klas reagują na informacje na ich temat - z uwzględnieniem źródła wiadomości oraz jej czasu trwania.
- Przeprowadzenie kompleksowej analizy efektywności prognoz instrumentów finansowych różnych klas za pomocą autorskiego podejścia LSTM-FENE i wykazanie ich wyższości w porównaniu z prognozami modeli ARMA-GARCH, ARMA-GARCH-FENE oraz LSTM.
- Stworzenie autorskich programów komputerowych w ogólnie dostępnych środowiskach programistycznych.

Ogólna ocena rozprawy

Rozprawa doktorska została przygotowana bardzo starannie pod względem językowym i edytorskim. Jest napisana jasno i przystępnie, co umożliwia potencjalny odbiór jej treści bardzo szerokiemu gronu czytelników. Badania ilościowe zostały dobrze zaplanowane – zaczynając od precyzyjnego opisu przygotowania danych, poprzez staranny opis wyników (uzupełniony Anekssem) aż po publikację kodów wykorzystanych przez Autora, co koresponduje z postulatem replikowalności badań naukowych. Na szczególne podkreślenie zasługuje dyskusja przedstawiona w rozdziale siódmym, dotycząca ograniczeń badania i możliwości jego kontynuacji.

Kwestie dyskusyjne i pytania do Autora

1. **Cel pracy i hipotezy badawcze.** Cel pracy został sformułowany jako *zbadanie możliwości zastosowania najnowszych metod przetwarzania języka naturalnego do zautomatyzowanej analizy wiadomości oraz postów o tematyce finansowej, gospodarczej i politycznej, a także uwzględnienia ich w finansowych modelach prognostycznych*. Wydaje mi się, że celem rozprawy było raczej zbadanie możliwości poprawy prognoz cen instrumentów finansowych poprzez uwzględnienie informacji zawartej w wiadomościach oraz postach o tematyce finansowej, gospodarczej i politycznej, która to informacja została wyekstrahowana metodami przetwarzania języka naturalnego. W sformułowanych przez Autora hipotezach brakuje określenia, w jaki sposób Autor będzie rozumiał efektywność. Brak tej definicji powoduje, że hipoteza zawsze może być zweryfikowana pozytywnie. Ostatecznie, formułując hipotezy 3 i 5, Autor nie wyjaśnia, czym są "modele wykorzystujące klasyczne metody sztucznej inteligencji" oraz "obecnie stosowane metody prognozowania szeregów czasowych". Obecnie stosuje się całe spektrum metod, łącznie z uwzględnieniem informacji płynącej z wiadomości i postów. Większość tych wątpliwości została jednak przez Autora wyjaśniona w rozdziale czwartym, gdzie jeszcze raz przytacza hipotezy pracy.



2. Rozdział pierwszy:

- Na stronie 25 Autor opisuje rynek kryptowalut i jako jedną z jego zalet wymienia z *góry określoną podaż pieniądza*. Jest to według mnie zbyt duży skrót myślowy, sugerujący, że kryptowaluty są rodzajem pieniądza. Dobrze byłoby ten wątek rozwinąć, powołując się na przykład na publikację (Marszałek, 2019).
- Podrozdział **1.4. Ekonometria finansowa a informacja na rynkach finansowych** – jako kluczowy z punktu widzenia celu i hipotez rozprawy – mógłby być bardziej rozbudowany. W przytaczanej literaturze brakuje na przykład istotnej pozycji na polskim rynku: Będowska-Sójka (2014).
- Na stronie 31 Autor przedstawia wzór 1.1 jako hipotezę rynku efektywnego. Brakuje źródła tego wzoru (w różnych publikacjach można spotkać różne wersje), a także wyjaśnienia, w jaki sposób liczona jest stopa zwrotu $R_{j,t+1}$.

3. Rozdział drugi:

- Modele klasy ARMA-GARCH zostały zasadniczo poprawnie opisane, natomiast brakuje wyraźnego połączenia części liniowej (ARMA) i nieliniowej (GARCH). Wydaje mi się, że opis zyskałby na czytelności, gdyby Autor zastosował w obu przypadkach te same symbole. W opisie modeli ARMA pojawia się symbol X_t , zamiast logarytmicznego zwrotu r_t , który wykorzystywany jest w opisie modelu GARCH (2.14). Dodatkowo, w modelu (2.14) Autor definiuje $r_t = \mu_t + \varepsilon_t$ - przy czym nie wyjaśnia, że μ_t jest to warunkowa (względem informacji posiadanej do chwili $t - 1$ włącznie) wartość oczekiwana, która w modelu ARMA-GARCH modelowana jest właśnie jednym z przedstawionych wcześniej modeli liniowych.
- Dodatkowo, sam symbol r_t jest dwójako zdefiniowany. na stronie 49 Autor pisze, że logarytmiczną stopę zwrotu liczy jako $\ln(r_t)$, a w późniejszej części rozprawy samo r_t oznacza logarytmiczną stopę zwrotu (por. wzór 2.28).
- Na stronie 53, definiując ε_t , Autor stwierdza, że jest to ciąg niezależnych zmiennych losowych o tym samym rozkładzie, ze średnią zero i zmienną wariancją σ_t^2 . Sugerowałabym uściślenie, że jest to wariancja warunkowa, bo obecne stwierdzenie sugeruje, że proces jest niestacjonarny. Dodatkowo, w dalszym ciągu opisu Autor zmienia symbol wariancji ε_t z σ_t^2 na h_t^2 (wzór 2.15).
- W tym samym wzorze 2.15 pojawia się symbol ψ_t , który nigdzie nie jest zdefiniowany.
- Wzór (2.19) może wywołać pewną konsternację u czytelnika. W (2.15) ε_t jest resztą z równania 2.14, więc z definicji jego wartość oczekiwana powinna wynosić 0. Dodatkowo, jego warunkowa wariancja podlegała procesowi (G)ARCH - natomiast w równaniu (2.19) jest stała. Wydaje mi się, że opis zaproponowany w pracy (Doman and Doman, 2009), na którą Autor również się powołuje, tj. dekompozycja reszty z równania opisującego zwrot $r_t = \mu_t + y_t$, $y_t = \varepsilon_t \sqrt{h_t}$, gdzie konsekwentnie ε_t jest białym szumem (lub ogólniej –iid o zerowej średniej i jednostkowej wariancji), jest bardziej czytelna.

- We wzorach 2.19 i 2.21 Autor podaje jako jeden z parametrów rozkładu wariancję (σ^2), podczas gdy we wcześniejszych wykorzystywał odchylenie standardowe (σ). Oczywiście, oba podejścia są poprawne, ale dobrze byłoby zdecydować się na jedno i stosować je konsekwentnie.
- Na stronie 57 Autor pisze, że nie ma badań, które pokazują, jak długi szereg należy wykorzystać do estymacji modelu GARCH. Nie jest to prawda. Przykładem takiej pracy jest choćby Hwang and Valls Pereira (2006).

4. Rozdział trzeci:

- Błędne odwołanie do publikacji [116] na stronie 92. Przy okazji, skoro Autor pisał rozprawę w LaTeXu, mógł wybrać wygodniejszy sposób cytowania (nazwisko, rok). W ten sposób uniknąłby podobnych pomyłek.
- Sformułowanie *analiza sentymentu* jest kalką językową. Trafniej byłoby użyć zwrotu *Analiza wydźwięku* (w kontekście słowa pisanego) lub *Analiza nastrojów* (w odniesieniu do inwestorów).

5. Rozdział czwarty:

- Wydaje mi się, że umieszczenie opisu danych w rozdziale przedstawiającym metodykę jest przedwczesne i bardziej pasuje do rozdziału piątego.
- Na stronie 120, na początku podrozdziału 4.4.1 pojawia się nieścisłość, która sugeruje, że modele ARMA-GARCH są modelami liniowymi. Do modeli liniowych zaliczamy modele typu ARMA, natomiast modele GARCH uważane są za nieliniowe (modelują zależności kwadratowe lub innego typu, jak np. model APARCH)
- Autor pisze, że analizy reszt modelu dokonuje za pomocą funkcji ACF i PACF. Dlaczego nie za pomocą testów statystycznych, np. Pearsona czy Ljung-Boxa? Ponadto, jeśli w grupie modeli wybranych przez Autora pojawiają się modele GJR, to warto też wspomnieć o testach obciążenia znakiem.
- Na stronie 103 Autor definiuje zwrot standaryzowany jako $\bar{r}_{t,n}$, przy czym po prawej stronie równania brakuje n . Domyślam, się, że w liczniku powinno być $r_{t,n}$ zamiast r_t ? Dodatkowo, nie zdefiniowano, czym jest n .
- We wzorze 4.2 na stronie 103 pod operatorem sumowania powinno być $\bar{r}_{t,n}^2$ - brakuje kwadratu oraz wyjaśnienia, czym jest T .
- W podrozdziale 4.5.3 Autor wyjaśnia, które błędy prognoz będzie brał pod uwagę przy ocenie modeli. Wydaje mi się, że ciekawie byłoby też wziąć pod uwagę testy, które pokazują, czy dobrze zaprognozowano sam kierunek zmian (wzrost lub spadek) - zwłaszcza, że w badaniu uwzględnia się wydźwięk informacji.

Czy podejście Autora pozwala na zweryfikowanie, w jaki sposób oddziałują na ceny informacje o wydźwięku negatywnym, a w jaki - pozytywnym? Pozwoliłoby to na zweryfikowanie

istnienia efektu dźwigni zdefiniowanego po raz pierwszy w publikacji (Black, 1976), tj. czy informacje o wydźwignię negatywnym rzeczywiście powodują większą reakcję inwestorów, niż informacje o wydźwignię pozytywnym.

6. Rozdział piąty:

- Na stronie 142 Autor pisze, że postać modelu ARMA-GARCH była wybierana na podstawie kryteriów informacyjnych (jeśli wskazania BIC i AIC się różniły, to podążano za AIC). Czy w takim razie pominięte zostało testowanie reszt, o którym Autor wcześniej wspominał?
- W podrozdziale 5.2 Autor przedstawia kryterium wyboru lepszego modelu - jako procentową zmianę błędu prognozy w stosunku do błędu modelu bazowego. Czy Autor rozważał zastosowanie innego podejścia - tj. przeprowadzenia testu statystycznej istotności różnic kolejnych błędów poprzez porównywanie prognoz parami i dopiero potem uśrednienie różnicy? W ten sposób można ocenić, czy szereg różnic błędów jest statystycznie różny od zera i byłaby to dodatkowa informacja, uzupełniająca test Diebolda-Mariano.
- Co oznacza wartość 0.25 jako dzienna liczba wiadomości? Czy jest to wartość uśredniona i oznacza, że dla danej spółki pojawiała się jedna informacja na 4 dni? Jeśli tak, to czy próbkowanie informacji śróddzienne (co 5 minut) nie jest nadmiarowe? Czy jeśli testy wskazują na wyższość modelu LSTM-FENE ponad ARMA-GARCH to nie jest to wynikiem wyższości modelu LSTM nad modelem ARMA-GARCH (co Autor wykazuje w kolejnym rozdziale) i może w przypadku instrumentów, co do których informacja pojawia się rzadko, wystarczająco dobry będzie model LSTM?
- Poproszę o wyjaśnienie, dlaczego Autor wybrał modele ARMA-GARCH do stawiania prognoz punktowych? Część GARCH odpowiada za modelowanie zmienności i jest wykorzystywana do stawiania prognoz zmienności lub uzupełnienia prognoz punktowych o przedział dopuszczalnych odchyleń. Czy nie wystarczyłyby same modele ARMA?
- Dodatkowo, interesujące byłoby porównanie prognoz z autorskiego modelu ze zwykłą prognozą naiwną.
- na stronie 165 Autor pisze, że "Prowadząc badania na danych zgromadzonych dla 5-minutowych interwałów udało się dodatkowo potwierdzić hipotezę, że *Wykorzystanie szeregu czasowego notowań o 5-minutowej częstotliwości jest wystarczające, aby zaobserwować zmiany wielkości stóp zwrotu analizowanych aktywów finansowych spowodowane publikacją wiadomości.* Wydaje mi się, że żeby potwierdzić tę hipotezę, należałoby przeprowadzić badania na danych o niższej i wyższej częstotliwości i określić, w którym momencie zmiany przestają być widoczne. Dodatkowo, na stronie 167 Autor dodaje uwagę: *Rynki kryptowalut i metali szlachetnych są wyjątkowo płynne, dlatego wykorzystanie interwałów 5-minutowych może być niewystarczające (...)* - kwestionując swój wcześniejszy wniosek.
- Jak można (intuicyjnie) zinterpretować wynik przedstawiony np. w Tabeli 5.10, że model LSTM-FENE osiągnął lepsze wyniki pod względem miernika RMSE dla wszystkich badanych kryptowalut i metali szlachetnych, a pod względem MAE - gorsze?

7. Rozdział szósty:

- Autor wprowadza do modelu jako szereg danych egzogenicznych wektor wyjściowy z modułu FENE. W zależności od wyboru wielkości $d_{encoder}$ jest to szereg 4-, 8- lub 16-wymiarowy. Nawet przy długości szeregu czasowego 4914 obserwacji (strona 193), jest to dość dużo jak na model GARCH. W związku z tym wskazane byłoby sprawdzenie stabilności parametrów z wykorzystaniem testu Nybloma. Może można wybrać najlepszą wartość $d_{encoder}$ korzystając z kryterium BIC, które preferuje modele oszczędne? Czy $d_{encoder}$ ma interpretację - n.p. czy jest to liczba postów, w których pojawia się wzmianka o instrumencie? I czy same wartości $d_{encoder}$ są interpretowalne?
- Autor pisze, że jeśli w danym przedziale pięciominutowym nie pojawiła się informacja, to wartości w $d_{encoder}$ są wyliczane na podstawie pustego ciągu znaków. Czy to oznacza, że w takim przypadku wektor zmiennych objaśniających przekazywanych do modelu ARMA-GARCH zawiera też jakieś wartości? Czy można je zinterpretować?
- W związku z tym, że modele ARMA-GARCH w większości przypadków okazały się lepsze bądź gorsze od modeli ARMA-GARCH-FENE, w celu wykazania wyższości modelu LSTM-FENE wskazane byłoby w tabelach 6.16-6.21 przypomnienie porównania modelu autorskiego ze "zwykłym" modelem ARMA-GARCH przynajmniej dla tych instrumentów, dla których okazał się statystycznie lepszy.
- Autor pisze o tym, że specyfikacja modelu bazowego oszacowanego na oknie o długości trzech miesięcy była utrzymywana we wszystkich kolejnych oknach czasowych. Jaki jest według doświadczenia Autora zalecany czas utrzymywania specyfikacji, a kiedy należałoby ją zmienić? Analizowany w rozprawie okres obejmuje również moment wybuchu wojny w Ukrainie - czy miało to wpływ na błędy prognoz?

Dodatkowe pytania do Autora:

W podrozdziale 5.3. Autor weryfikuje wpływ źródła informacji na poprawę zdolności progностycznych modelu. Czy według wiedzy i doświadczenia Doktoranta miałoby sens przeprowadzenie osobnej analizy dla kryptowalut, z pominięciem wiadomości z portali społecznościowych?

Źródłem danych wykorzystywanych w pracy jest serwis AlphaVintage. Czy Autor mógłby wyjaśnić, w jaki sposób serwis podaje dane dotyczące kryptowalut? Czy są to dane z jednej konkretnej giełdy, czy uśrednione z kilku wiodących? Czy były to giełdy scentralizowane, czy zdecentralizowane? Jeśli są to dane uśrednione, to może to wyjaśniać, dlaczego rozwiązanie Autora nie było lepsze w tym przypadku, niż rozwiązania klasyczne. Czy miałoby sens przeprowadzenie badania na konkretnej giełdzie (np. o największej płynności) i zweryfikowanie wyników? Czy możemy się spodziewać lepszych wyników na giełdach scentralizowanych, czy zdecentralizowanych – czy nie powinno mieć to znaczenia?

Czy Autor mógłby intuicyjnie wyjaśnić, w jaki sposób jego model ocenia wydźwięk informacji i w jaki sposób jego podejście różni się od podejścia "słownikowego", w którym wykorzystywane są słowniki o określonej tematyce (Apel and Grimaldi, 2014; Bennani and Neuenkirch, 2017; Apel et al., 2022) i

wydzwięk informacji bada się poprzez zliczanie zawartych w niej słów o danym wydzwięku (podanym przez słownik)? W tabelach wynikowych rozprawy jest podana liczba słów, ale bez podziału na wydzwięk pozytywny i negatywny.

Konkluzja

Chciałabym wyraźnie podkreślić, że przedstawione w recenzji uwagi nie wpływają znacząco na moją ogólną ocenę tej bardzo dobrej rozprawy doktorskiej. Uważam, że dysertacja pana magistra Piotra Bednarskiego jest pracą bardzo dobrze napisaną. Autor podjął się istotnego problemu prognozowania cen instrumentów finansowych z wykorzystaniem informacji tekstowej próbkowanej z pięciominutową częstotliwością. Poprawnie zidentyfikował lukę badawczą, postawił hipotezy i zaprojektował badanie umożliwiające ich weryfikację. Podjęty temat jest aktualny i wpisuje się w główny nurt badań w zakresie finansów empirycznych. Recenzowana rozprawa doktorska stanowi istotny wkład w rozwój dziedziny ekonomia i finanse, pokazując, że uwzględnienie w modelu prognostycznym zaawansowanych technologii przetwarzania języka naturalnego może przyczynić się do istotnej poprawy prognoz. W związku z tym, proponowane przez Autora rozwiązanie ma też wymierne zastosowanie praktyczne. Dojrzałości naukowej Doktoranta dowodzi fakt, iż ma on świadomość również ograniczeń swojego podejścia, które przedstawia w końcowej części rozprawy, proponując jednocześnie kierunki jego rozwinięcia i dalszych badań.

W mojej ocenie przedstawiona rozprawa potwierdza zdolność Doktoranta do prowadzenia samodzielnych badań naukowych, jego wiedzę teoretyczną na temat przedstawianych zagadnień, stanowi wkład w rozwój nauki polskiej, a także przedstawia oryginalne rozwiązanie konkretnego problemu badawczego.

Podsumowując przedstawione oceny i uwagi, stwierdzam, że recenzowana rozprawa doktorska spełnia warunki określone w ustawie z dnia 20 lipca 2018 roku, *Prawo o szkolnictwie wyższym i nauce* (Dz. U. 2023 poz. 742 z późn. zm.). W związku z tym wnoszę o dopuszczenie pana mgr. Piotra Bednarskiego do dalszych etapów przewodu doktorskiego, prowadzących do nadania mu stopnia doktora nauk ekonomicznych w dyscyplinie ekonomia i finanse.

Agata Kliber

Literatura

- Apel, M. and Grimaldi, M. B. (2014). How informative are central bank minutes? *Review of Economics*, 65(1):53–76.
- Apel, M., Grimaldi, M. B., and Hull, I. (2022). How much information do monetary policy committees disclose? Evidence from the FOMC's minutes and transcripts. *Journal of Money, Credit and Banking*, (54):1459–1490.

- Będowska-Sójka, B. (2014). *Wpływ informacji na ceny instrumentów finansowych: analiza danych śróddziennych*. Wydawnictwo Uniwersytetu Ekonomicznego.
- Bennani, H. and Neuenkirch, M. (2017). The (home) bias of European central bankers: New evidence based on speeches. *Applied Economics*, 49(11):1114–1131.
- Black, F. (1976). Studies of stock market volatility changes. *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics Section*, pages 177–181.
- Doman, M. and Doman, R. (2009). *Modelowanie zmienności i ryzyka. Metody ekonometrii finansowej*. Wolters Kluwer Polska SA.
- Hwang, S. and Valls Pereira, P. L. (2006). Small sample properties of GARCH estimates and persistence. *The European Journal of Finance*, 12(6–7):473–494.
- Marszałek, P. (2019). Kryptowaluty–pojęcie, cechy, kontrowersje. *Studia BAS*, (1):105–125.

