

Recenzja pracy doktorskiej Pani mgr Joanny Michalak pod tytułem *Modelowanie zależności między zwerbalizowanymi nastrojami i emocjami użytkowników serwisu społecznościowego Twitter a indeksami giełdowymi* napisanej pod kierunkiem Pana prof. dr hab. Tadeusza Kufła

Formalną podstawą do sporządzenia niniejszej recenzji stanowi pismo 11.WNEiZ/530-31/2023 Pani Profesor dr hab. Magdaleny Osińskiej, Przewodniczącej Rady Dyscypliny Naukowej Ekonomia i Finanse Uniwersytetu Mikołaja Kopernika w Toruniu, informujące o powołaniu mnie na recenzenta rozprawy doktorskiej Pani mgr Joanny Michalak pod tytułem *Modelowanie zależności między zwerbalizowanymi nastrojami i emocjami użytkowników serwisu społecznościowego Twitter a indeksami giełdowymi* napisanej pod kierunkiem Pana prof. dr hab. Tadeusza Kufła.

W kolejnych częściach recenzji chciałbym przedstawić ocenę poszczególnych części rozprawy doktorskiej, następnie przedstawić uwagi dotyczące całości pracy i sformułować konkluzję.

1 Charakterystyka Wstępu

W początkowej części *Wstępu* Doktorantka wskazuje na znaczenie wpływu czynników behawioralnych na sposób podejmowania decyzji finansowych przez człowieka. Zdaniem Autorki, prace badawcze prowadzone w tym obszarze nie pozwoliły jeszcze na zbadanie wszystkich aspektów tego zagadnienia, a istniejąca luka badawcza stanowiła główną przesłankę do podjęcia decyzji o ukierunkowaniu przygotowanej rozprawy doktorskiej na tę problematykę.

Głównym celem pracy jest *przeprowadzenie empirycznej weryfikacji zależności między nastrojami i emocjami eksplorowanymi z serwisu społecznościowego Twitter a indeksami giełdowymi*. Ten, dość ogólnie sformułowany cel, został doprecyzowany poprzez sformułowanie pięciu pytań badawczych i celów szczegółowych pracy. Zawarty we *Wstępie* zestaw celów i pytań badawczych w zasadzie odpowiada treściom znajdującym się w dalszej części rozprawy. Aczkolwiek brakuje mi w nich mocniejszego wypuklenia problematyki emocji omawianej szczegółowo w rozdziale drugim i stanowiących istotny czynnik przy projektowaniu i przeprowadzaniu procesu badawczego.

W tej części rozprawy przedstawione zostały dwie hipotezy, weryfikacji których mają służyć rozważania zamieszczone w pracy. Pierwsza z nich mówi o istnieniu istotnej statystycznie współzależności pomiędzy aktywnością użytkowników serwisu społecznościowego Twitter a wolumenem obrotu giełdowego. Zaś druga głosi, że wskaźniki nastroju są dobrym predyktorem wolumenu obrotu giełdowego, a wskaźniki emocji są efektywniejszym predyktorem tego wolumenu. Hipotezy te zostały opatrzone przez Doktorantkę komentarzem zawierającym odwołania do wcześniejszych badań. Zastanawiam się, dlaczego Autorka powołuje się w tej części tekstu na prace z 2004 i 2011 roku. Czy problematyka ta nie była poruszana w kolejnych latach? Czy obecny poziom wykorzystania serwisu Twitter przez inwestorów jest porównywalny z sytuacją z 2011 roku?

Dalsza część *Wstępu* zawiera omówienie zawartości kolejnych rozdziałów pracy.

2 Charakterystyka Rozdziału I

Tematyka *Rozdziału I* dotyczy serwisów społecznościowych i powstających w nich danych.

Autorka rozpoczyna swoje rozważania od przedstawienia definicji i klasyfikacji mediów społecznościowych (temu zagadnieniu poświęcony jest podrozdział 1.1 *Definicja i klasyfikacja mediów społecznościowych*). W tej części pracy znajdujemy szereg informacji porządkujących dalsze rozważania i pozwalających na umiejscowienie prowadzonych prac badawczych w szerokim obszarze komunikacji internetowej.

Podrozdział 1.2 nosi tytuł *Pojęcie i metody badawcze big social data* i jest poświęcony problematyce *big social data*. Tytuł podrozdziału zapowiada, że prezentowane tu treści dotyczyć będą samego pojęcia oraz metod analizy tego typu danych. O ile samo pojęcie zostało przedstawione szeroko i wraz z powołaniem na publikacje dotyczące tego zakresu, to omówienie metod właściwych do analizy tego typu zasobów jest raczej powierzchowne. Zdaję sobie sprawę, że dokładne zaprezentowanie tej tematyki byłoby bardzo praco- i czasochłonne. Dlatego może należało nie sygnalizować kwestii metod analizy w tytule podrozdziału – w ten sposób oczekiwanie dotyczące omówienia tej kwestii pozostałoby bezzasadne.

W podrozdziale 1.3 *Twitter jako źródło informacji oraz wybrane obszary jego wykorzystania* zaprezentowany został Twitter, a w szczególności sposób funkcjonowania serwisu i struktura wiadomości.

Kolejny podrozdział to 1.4 *Reprezentatywność użytkowników Twittera w populacji*. Poświęcony jest on pobieraniu danych przesyłanych w ramach serwisu Twitter. Po krótkim omówieniu mechanizmów pozwalających na pozyskanie pewnej części komunikatów, Doktorantka skupiła się na problemie reprezentatywności i omawia kolejno sposób pobierania próbki, wpływ mechanizmu przekazywania wiadomości na reprezentatywność czy brak możliwości pobierania komunikatów przesyłanych przez określoną grupę użytkowników. Wnioski płynące z przedstawionej analizy wskazują na brak argumentów potwierdzających reprezentatywność danych udostępnianych przez platformę Twitter.

3 Charakterystyka Rozdziału II

Tematyka rozdziału II *Stany afektywne i ich rozumienie i konsekwencje* jest zgodna z treścią zapowiadaną przez tego tytuł.

Doktorantka słusznie zauważa, że przygotowana rozprawa doktorska ma charakter interdyscyplinarny, łączy problematykę psychologii, medycyny, ekonomii czy finansów. I z tego powodu konieczne jest dokładne zdefiniowanie pojęć wykorzystywanych w badaniach. I taką rolę ma, zgodnie z koncepcją Autorki, odgrywać rozdział II. Rozdział ten prezentuje psychologiczno-neurologiczne podstawy prowadzonych badań. Składa się z siedmiu części (część wprowadzająca i podrozdziały numerowa od 2.1 do 2.6).

Pierwszą częścią rozdziału ma charakter wprowadzający i poprzedza punkt 2.1. W tej części tekstu Autorka sygnalizuje zagadnienia przedstawiane w kolejnych podrozdziałach. W mojej ocenie jest to najślabza część rozdziału II. Definicja emocji wymaga dopracowania (na stronie 35 znajdujemy stwierdzenie *emocja jest empiryczną hipotezą dotyczącą natury emocji bądź składu jej klas*). Warto

byłoby podać kilka innych definicji (Autorka wskazuje na ich wielość, ale ich nie przytacza w uporządkowany sposób). Kolejne partie tekstu sygnalizują wybrane zagadnienia dokładnie omawiane w podrozdziałach 2.1 – 2.6. O ile prezentacja zawarta w kolejnych podrozdziałach ma charakter uporządkowany, to w tej części dość mocno odbiega od takiego charakteru. Jestem przekonany, że rozdział II zyskał bardzo na jakości, gdyby część wprowadzająca została skompresowana do kilku zdań.

W odróżnieniu od mojej oceny części wprowadzającej, podrozdziały 2.1 – 2.6 oceniam w pełni pozytywnie. Przede wszystkim z uwagi na włączenie tej problematyki do zakresu pracy doktorskiej. Zaprezentowane tu rozważania stanowią bardzo istotną podbudowę prowadzonych badań. Dzięki nim konstruowane w dalszej części modele oparte są nie tylko na zidentyfikowanych automatycznie zależnościach statystycznych pomiędzy liczbowymi charakterystykami tekstu, ale opierają się na wiedzy z zakresu neurologii czy psychologii. Lektura tego rozdziału potwierdza bardzo wysoki poziom wiedzy Doktorantki.

Podrozdział 2.1. nosi tytuł *Emocje w klasycznych teoriach psychologicznych*. W tej części swojej pracy Autorka omawia takie zagadnienia jak: cechy emocji, sposoby ich wyrażania i poznawanie, wpływ na myślenie, funkcjonowanie pamięci i funkcje poznawcze człowieka oraz pochodzenie emocji. Treści zaprezentowane w omawianym podrozdziale jednoznacznie wskazują, że emocje związane są z biologiczną sferą człowieka i stanowią bardzo istotny czynnik wpływający na jego funkcjonowanie.

Podrozdziałowi 2.2. Doktorantka nadała tytuł: *Podstawowość emocji i ich klasyfikacja w ujęciu klasycznym*. Punktem wyjścia do przedstawionych tu rozważań jest wspomniany w końcowej części poprzedzającego podrozdziału podział emocji na pierwotne i wtórne. Za dokonaniem takiego podziału przemawia zarówno niepodzielność emocji podstawowych i złożoność emocji wtórnych jak i różny sposób ich powstawania i przetwarzania przez układ nerwowy człowieka. Bardzo dobrym podsumowaniem rozważań zawartych w tym podrozdziale jest Tabela 6, która zawiera zestawienie emocji podstawowych w sześciu różnych teoriach dotyczących tego zagadnienia. Autorka wskazuje na różne podstawy teoretyczne poszczególnych teorii zajmujących się naturą emocji. Wspomina o możliwości ich postrzegania z perspektywy biologicznej, socjo-kulturowej czy ewolucyjnej.

W podrozdziale 2.3. *Źródła wyzwalania emocji w ujęciu klasycznym* możemy zapoznać się z prezentacją poglądów różnych badaczy na źródła emocji. Można tu wyodrębnić kilka różnych punktów widzenia. Przedstawiciele pierwszego z nich twierdzą, że emocje powstają w odpowiedzi na bodźce. Inna grupa badaczy w powstawaniu emocji dostrzega również rolę procesów poznawczych człowieka. Są również tacy, którzy są przekonani, że emocje rodzą się wskutek relacji zachodzących pomiędzy człowiekiem a otoczeniem.

Z kolei podrozdział 2.4. *Emocje w ujęciu neuropsychologii* pozwala spojrzeć na emocje z punktu widzenia układu nerwowego człowieka. Takie spojrzenie zmienia sposób badania emocji. Również w inny sposób rozpatrywane jest w tym aspekcie oddziaływanie emocji na funkcjonowanie człowieka, w tym na podejmowane przez niego decyzje inwestycyjne.

W podrozdziale 2.5 *Emocje a nastrój* Doktorantka wskazuje na potrzebę rozróżnienia emocji i nastroju. Te dwa stany różnią się między sobą czasem istnienia – emocje trwają przez stosunkowo krótki okres, zaś nastrój utrzymuje się znacznie dłużej. W przypadku emocji zwykle można wskazać obiekt, z którym są one powiązane. Jest to znacznie trudniejsze w przypadku nastroju, gdzie relacja ta ma znacznie bardziej rozmyty charakter. Emocje i nastrój różnią się swoją dynamiką, która jest znacznie większa w przypadku emocji. Pomimo różnic, emocje i nastrój są mocno ze sobą powiązane i wzajemnie oddziałują na siebie. Oba te zjawiska mają wpływ na decyzje inwestorów. Interesującą kwestią jest również badanie emocji w oparciu o komunikaty tekstowe przesyłane na platformie Twitter. W tym kontekście pojawia się wiele pytań badawczych: czy emocje, jako stany krótkotrwałe i dynamicznie się zmieniające, można

badać w oparciu o stosunkowo rzadko wysyłane komunikaty tekstowe? w jakim stopniu mechanizm przekazywania komunikatów wpływa na obraz emocji inwestorów? Sprowadza się to do zagadnienia związanego z definicją obiektu badań – czy badaniu podlegają emocje indywidualnych osób, czy też emocje zagregowane na poziomie grup użytkowników.

Wspomnianemu powyżej zagadnieniu emocji grupowych poświęcony jest podrozdział 2.6. noszący tytuł *Procesy grupowe i mechanizm zarażania afektywnego*. W tej części pracy omówione zostało zagadnienie transmisji emocji w sieci społecznościowej przyjmujące postać procesu zarażania afektywnego. Autorka wskazuje, że ten rodzaj zarażania dotyczy zarówno emocji jak i nastroju i prowadzi do powstania afektów kolektywnych (ich przykładem była „arabska wiosna”).

Rozdział II oceniam bardzo wysoko. Stanowi interesujący i wartościowy punkt wyjścia do dalszych badań nad wydźwiękiem wiadomości przesyłanych na platformie Twitter. Tego typu rozważań brakuje w polskiej literaturze. Jest to bardzo istotny wkład w rozwój ekonomii behawioralnej. Zdaję sobie sprawę, że operacjonalizacja wszystkich przedstawionych tu idei, w szczególności do analizy tekstów polskojęzycznych, może być zadaniem trudnym i czasochłonnym.

4 Charakterystyka rozdziału III

Rozdział III został zatytułowany *Proces badawczy dla modelowania zależności między Twitter a giełdą*. Zawiera omówienie sposobu przygotowania danych i metod badawczych stosowanych w dalszych analizach. Rozdział ten składa się z dwóch podrozdziałów.

Podrozdział 3.1. *Eksploracja danych tekstowych* poświęcony jest prezentacji metod służących do analizy danych tekstowych (ze szczególnym uwzględnieniem komunikatów przesyłanych za pośrednictwem platformy Twitter).

W początkowej części tego podrozdziału (stanowiącej punkt 3.1.1) przedstawiona została specyfika wiadomości przesyłanej za pomocą tej platformy. Zwrócono uwagę na takie zagadnienia jak ich ograniczona długość i wynikające z tego faktu związanie z jednym tematem, występowanie znaków i symboli specjalnych (hashtagi, cashtagi, emotikony), nieformalny charakter wiadomości (skrótły, określenia slangowe), zwrócono uwagę na występowanie błędów ortograficznych oraz na wielojęzyczność wiadomości. Doktorantka zwraca również uwagę na mechanizm przekazywania tweetów, możliwość osadzania treści multimedialnych i możliwość generowania komunikatów przez boty. Wszystkie wymienione elementy powinny być uwzględnione w trakcie analizy. W wielu przypadkach zaprojektowanie i wdrożenie właściwych rozwiązań może stanowić duże wyzwanie.

Natomiast punkt 3.1.2 prezentuje dwa podejścia, które doktorantka zdecydowała się zastosować w ocenie sentymentu wiadomości tekstowych. W pierwszym podejściu wykorzystano uczenie maszynowe, zaś drugie podejście ma charakter słownikowy.

Budując system oceny sentymentu wykorzystujące nadzorowane uczenie maszynowe założono, że tweety poddawane są czyszczeniu, następnie przekształcane do macierzy TFDIF i przypisywane przez klasyfikator je jednej z dwóch klas (nacechowanych pozytywnie lub nacechowanych negatywnie). Model budowany był w oparciu o zbiór uczący, w którym etykiety opisujące nacechowanie przypisane zostały automatycznie. Tak więc skonstruowany klasyfikator jest aproksymatorem zasad zastosowanych do klasyfikacji zbioru uczącego. W takiej sytuacji pojawia się pytanie, czy nie lepiej byłoby zastosować oryginalne reguły zastosowane do przypisania etykiet, a nie je przybliżać zastosowanym klasyfikatorem.

Drugie podejście wykorzystuje słowniki pozwalające na identyfikację wyrazów i fraz opisujących emocje lub wydzwięk. Wykorzystano dwa słowniki: VADER i NRC Emotion Lexicon.

Podrozdział 3.2. nosi tytuł *Ekonometryczna analiza związku między social media a giełdą*. W pierwszej części podrozdziału znajdujemy omówienie doniesień literaturowych dotyczących modelowania związków pomiędzy kształtowaniem się indeksów giełdowych a zawartością komunikatów przesyłanych za pomocą platformy Twitter. Następnie znajdujemy krótkie omówienie procedury badawczej zastosowanej w pracy. Zastosowano model wektoro-autoregresyjny wykorzystujący informacje o stanie giełdy oraz informacje o wydzwięku tweetów. Opis zastosowanych modeli ma dość ogólny charakter. Z uwagi na to, że wykorzystane informacje pochodzą z dwóch różnych systemów (Twitter i giełda), które generują dane z różną częstotliwością, może byłoby warto dokładniej opisać zastosowaną procedurę, która doprowadziła do porównywalności tych danych.

5 Charakterystyka rozdziału IV

Rozdział 4. nosi tytuł *Analiza zależności między wybranymi wskaźnikami Twittera a indeksami giełdowymi*. Rozdział ten prezentuje wyniki przeprowadzonych badań. Wstępna część rozdziału (poprzedzająca podrozdział 4.1) zawiera charakterystykę szeregów czasowych utworzonych na podstawie komunikatów przesyłanych na platformie Twitter. Doktorantka zwraca uwagę na takie elementy jak cykliczność, występowanie trendu, heteroskedastyczność oraz występowanie długiej pamięci procesów. Tutaj też przedstawiona jest krótka charakterystyka modeli VAR (wektorowej autoregresji), które zostały wykorzystywane w trakcie dalszych badań.

Podrozdział 4.1. zatytułowany został *Charakterystyka spółek, indeksu oraz agencji prasowej*. W moim odczuciu ten tytuł jest mylący, gdyż poświęcony jest charakterystyce tweetów dotyczących wspomnianych w tytule podmiotów, a nie samych podmiotów. Znajdujemy tu informację o zakresie prowadzonych badań. Zakres czasowy ograniczony był do lat 2016-2017. Badania dotyczyły notowań firm Apple, Amazon, Facebook, Google, Bloomberg oraz indeksu S&P500. Doktorantka wskazuje na uwzględnioną tematykę tweetów, która w przypadku firm została podzielona na 5 grup, zaś w przypadku dyskusji dotyczącej indeksu S&P500 – na 3 grupy. Nie znalazłem jednak wyjaśnienie dotyczącego sposobu dokonywania klasyfikacji do wyróżnionych grup tematycznych. Identyfikację emocji przeprowadzono w sposób opisany w rozdziale 3.1. Tabela 12 zawiera charakterystykę uzyskanych wyników. Zastanawiają mnie wyniki przedstawione w tabelach 13, 14 i 15. Zaprezentowano tu informacje o odsetku tweetów, w których zidentyfikowano elementy opisujące emocje lub nastrój za pomocą słownika NRC (tabela 13), słownika VADER (tabela 14) i skonstruowanego klasyfikatora (tabela 15). Wydaje się, że uzyskane wyniki różnią w dość dużym stopniu. Co jest przyczyną tak dużego zróżnicowania i które z uzyskanych mierników lepiej opisują rzeczywistość? W kolejnej części omawianego podrozdziału przedstawiono informacje o wolumenie obrotu giełdowego i wolumenie komunikatów w rozbiciu na analizowane podmioty.

Podrozdziałowi 4.2. został nadany tytuł *Wyniki analizy danych tekstowych*. Również w tym przypadku tytuł nie oddaje zawartości podrozdziału (dokładnie mówiąc jego zakres tematyczny jest znacznie szerszy od zawartości tej części pracy). Podrozdział ten poświęcony jest ocenie klasyfikatorów tworzonych przez Doktorantkę przy wykorzystaniu naiwnego klasyfikatora Bayesa i klasyfikatora SVM na podstawie zbioru danych *Senti140*. Moja podstawowa uwaga w tym zakresie została przedstawiona już powyżej – jest to model aproksymujący inny model, a nie rzeczywistą ocenę sentymentu. Sam sposób budowy modelu został przeprowadzony prawidłowo. Uzyskane wyniki w ujęciu bezwzględny nie są rewelacyjne, ale typowe dla tego typu zastosowań. Również dobór sposobu reprezentacji tekstu jest zgodny z oczekiwaniami dla tego typu podejścia. Podsumowując ten podrozdział można stwierdzić,

że Doktorantka wykazała się bardzo dobrym opanowaniem techniki budowy modeli klasyfikacyjnych opartych na macierzy częstości, zaś uzyskane wyniki są zgodne z oczekiwaniami.

Podrozdział 4.3. nosi tytuł *Analiza korelacji między wskaźnikami Twittera a giełdą*. W tym przypadku zawartość podrozdziału w pełni pokrywa się z brzmieniem tytułu. Autorka analizuje zależność korelacyjną pomiędzy analizowanymi w pracy wskaźnikami opisującymi stan giełdy i komunikację prowadzoną na platformie Twitter. Uzyskane wyniki są bardzo interesujące. Pierwszy ich aspekt, który mnie zainteresował jest związany z istnieniem pozytywnej korelacji pomiędzy obrotem giełdowym i wydawałoby się przeciwnymi znaczeniowo (negatywnymi i pozytywnymi) opiniami. Czy ten fakt można zinterpretować jako wskazujący na korelację aktywności (zarówno tej o pozytywnym, jak i negatywnym wydźwięku) na platformie Twitter a aktywnością giełdową (mierzoną wolumenem obrotu). Jeżeli Doktorantka zgadza się z takim stwierdzeniem (a po przeczytaniu tekstu podrozdziału, jestem przekonany, że tak), to czy zamiast zwykłego współczynnika korelacji nie należało liczyć korelacji cząstkowej pomijającej wpływ wolumenu obrotu? Uzyskane wartości współczynników korelacji nie świadczą o mocnym związku pomiędzy komunikacją na platformie Twitter a zachowaniem giełdy. W jakim stopniu na uzyskanie takich wyników wpływ miał rzeczywisty brak powiązań, a w jakim stopniu niedoskonałość narzędzi badających zawartość komunikatów? Zawartość podrozdziału 4.3 (podobnie jak w przypadku części 4.2 pracy) świadczy z jednej strony o dużej wiedzy i umiejętnościach praktycznych Doktorantki, ale z drugiej strony dowodzi konieczności prowadzenia dalszych badań w zakresie analizy korelacyjnej zjawisk zachodzących w świecie wirtualnym i świecie giełdowym.

Kolejny etap badań został przedstawiony w podrozdziale 4.4 zatytułowanym *Analiza krótkookresowej dynamiki z wykorzystaniem modelu VAR*. Zastosowanie modeli wektorowej autoregresji pozwoliło na jednoczesne ujęcie zależności pomiędzy wieloma zmiennymi uwzględnionymi w badaniach. Początkowa część podrozdziału zawiera opis procedury budowy modelu. Ten fragment pracy wskazuje na bardzo dobrą znajomość tej tematyki przez Doktorantkę. Znaczenie dyskusji prowadzonej w sieci Twitter została przeanalizowana między innymi poprzez odpowiedź modelu na impuls jednostkowy. Zasadniczą część podrozdziału stanowi prezentacja wyników badań dotyczących analizy przyczynowości pomiędzy dyskusją tweeterową a giełdą. Uzyskane wyniki potwierdzają wpływ dyskusji na zachowanie się giełdy w sensie przyczynowości Grangera. Znacznie mniej argumentów przemawia za istnieniem zależności odwrotnej. Autorka przeprowadziła też badanie kierunku zmian wolumenu obrotu w zależności od przebiegu dyskusji w sieci społecznościowej Twitter. Uzyskane wyniki potwierdzają wcześniejsze spostrzeżenia, że sam fakt dyskusji, a nie jej wydźwięk, wpływa na wzrost obrotów.

Podrozdział 4.5 zatytułowany jest *Podsumowanie*. Zawarte w nim rozważania są nie tylko podsumowaniem czwartego rozdziału, ale również całej pracy (osobiście uważam, że dla pracy byłoby lepiej, gdyby wnioski z całości zostały przedstawione w *Podsumowaniu* mającym charakter oddzielnego rozdziału, a nie podrozdziału). Doktorantka odniosła się do dwóch głównych hipotez pracy. Zdaniem Autorki uzyskane wyniki potwierdzają istnienie zależności statystycznej pomiędzy aktywnością inwestorów w ramach sieci Twitter a wolumenem obrotu. Natomiast istnienie związku pomiędzy wydźwiękiem wypowiedzi i wolumenem obrotu nie daje jednoznacznych wyników (badania pokazały jednak, że wskaźniki emocji są mocniej powiązane z wolumenem obrotu niż wskaźniki nastroju).

6 Całościowa ocena pracy

W tej części recenzji chciałbym sformułować kilka moich spostrzeżeń dotyczących całości rozprawy doktorskiej *Modelowanie zależności między zwerbalizowanymi nastrojami i emocjami użytkowników serwisu społecznościowego Twitter a indeksami giełdowymi* przygotowanej przez Panią mgr Joannę Michalak.

1. Tematyka pracy jest bardzo interesująca. Autorka zajmuje się badaniem zależności pomiędzy wskaźnikami opisującymi charakter komunikacji prowadzonej za pośrednictwem sieci społecznościowej Twitter a indeksami giełdowymi. Choć problematyka ta występowała już wcześniej w literaturze światowej, to kilka aspektów pracy ma charakter nowatorski. Poziom naukowy prowadzonych badań w pełni odpowiada wymogom stawianym kandydatom do uzyskania stopnia naukowego doktora.
2. Obszar badawczy pracy jest związany z finansami behawioralnymi. Stąd też ułożenie tej pracy w ramach dyscypliny *ekonomia i finanse* jest właściwe.
3. Praca ma w wielu miejscach charakter interdyscyplinarny. Jest to szczególnie wyraźnie widoczne w części dotyczącej badań zawartości komunikatów tekstowych przesyłanych za pośrednictwem platformy Twitter. Ich podstawy teoretyczne Autorka bardzo mocno usadowiła w obszarze psychologii i neurobiologii. Jest to dużą zaletą tej pracy. Szkoda jedynie, że to sprzężenie jest znacznie wyraźniej widoczne w części teoretycznej pracy niż w jej części empirycznej, w której wykorzystywana jest jedynie część wiedzy zaprezentowanej w rozdziale drugim rozprawy.
4. W części empirycznej pracy Doktorantka z powodzeniem łączy podejście eksploracyjne (eksploracyjna analiza dokumentów tekstowych mająca na celu ocenę ich wydźwięku) z podejściem ekonometrycznym (wykorzystanie wektorowej autoregresji).
5. Uważam, że hipotezy badawcze pracy zostały poprawnie sformułowane oraz proces ich weryfikacji został przeprowadzony właściwie. Dobór sposobu prowadzenia badań mających na celu weryfikację hipotez należy uznać za właściwy (pewne kwestie dyskusyjne dotyczące prowadzenia badań przedstawiłem we wcześniejszych punktach recenzji, traktuję je jednak jako punkt wyjścia do dyskusji w trakcie obrony dotyczącej szczegółów procesu badawczego, a nie jako elementy negujące sposób prowadzenia badań, etapy procesu badawczego czy wybór metod i narzędzi badawczych).
6. Bardzo wysoko oceniam stosowany przez Doktorantkę sposób interpretacji uzyskanych wyników. W wielu miejscach Autorka zauważa niejednoznaczności związane z uzyskanymi wynikami, wskazuje na słabe strony zastosowanego podejścia, z dużą ostrożnością podchodzi do formułowania kategoriycznych sądów. Wszystkie te cechy wskazują na bardzo dużą wiedzę Doktorantki i doświadczenie w prowadzeniu badań empirycznych.
7. Doktorantka potrafi wyraźnie rozgraniczyć kwestie będące przedmiotem badań od zagadnień, które nie zostały w pełni uwzględnione w procesie badawczym i które mogą stanowić temat przyszłych prac Autorki.
8. Praca oparta jest na właściwie dobranej literaturze.
9. Recenzowany egzemplarz pracy liczył 261 stron, przy czym 161 stron liczy zasadnicza część rozprawy, a kolejne sto stron zawiera przede wszystkim bardzo dokładnie udokumentowane wyniki badań.
10. Praca napisana jest poprawnym językiem. Jej tekst został przygotowany bardzo starannie. Prezentacja wyników ma postać tabelaryczną lub graficzną. Elementy te zostały bardzo starannie przygotowane i opisane.

7 Konkluzja

Po zapoznaniu się z rozprawą doktorską *Modelowanie zależności między zwerbalizowanymi nastrojami i emocjami użytkowników serwisu społecznościowego Twitter a indeksami giełdowymi* przygotowaną przez Panią mgr Joannę Michalak pod kierunkiem prof. dr hab. Tadeusza Kufla stwierdzam, że praca stanowi oryginalne rozwiązanie problemu naukowego, zaś Doktorantka posiada wiedzę teoretyczną w dyscyplinie *ekonomia i finanse* oraz umiejętność prowadzenia pracy naukowej. W związku z tym

stwierdzam, że praca spełnia wymogi ustawowe zawarte w art. 187 Ustawy z dnia 20 lipca 2018 roku
Prawo o Szkolnictwie Wyższym i Nauce i wnoszę o jej dopuszczenie do publicznej obrony.

Podpisane elektronicznie przez
Paweł Lula (Certyfikat
kwalifikowany) w dniu
2023-08-31.

P. Lula