

**Recenzja dysertacji doktorskiej mgra Raphaela Korbiniana Reinwalda, pt. 'On a Comparative Analysis of Industrial Credit Portfolio Risk Models Versus a New Support Vector Machine – Based Approach', napisanej na Wydziale Ekonomii Uniwersytetu Gdańskiego, pod kierunkiem naukowym prof. UG dr. hab. Michała Suchanka**

1. Uwagi wstępne

Deklarowanymi dziedziną i dyscypliną badań przeprowadzonych w pracy są odpowiednio nauki społeczne i ekonomia i finanse. Lektura dysertacji pozwala stwierdzić, że - zarówno w odniesieniu do jej części teoretycznej jak i empirycznej - przyporządkowanie to jest prawidłowe, czym autor wypełnia wymogi formalne. Można dodać, iż za spełnione wręcz z nadstatkiem należy uznać również oczekiwania względem wniesienia przez omawianą pracę wartości dodanej do aktualnego stanu wiedzy naukowej.

Bardziej szczegółowo określając obszar, w którym sytuują się przeprowadzone w pracy badania można stwierdzić, że stanowi go analiza ryzyka kredytowego i ocena skuteczności modeli pozwalających to ryzyko szacować. W tym względzie autor wykazał się dogłębną znajomością tematyki i to zarówno w odniesieniu do jej podstaw teoretycznych, jak też metod oceny praktycznych zastosowań, tworzonych w oparciu o te teoretyczne podstawy, modeli. Na uwagę zasługuje także kompleksowa znajomość regulacji prawnych obowiązujących podmioty funkcjonujące w sferze działalności kredytowej oraz literatury przedmiotu.

W odniesieniu do wyboru samego tematu, co zostanie omówione bardziej szczegółowo poniżej, należy przyznać, że wybór ów, zarówno w perspektywie znaczenia rynków finansowych dla współczesnej gospodarki, jak też zmian w ich funkcjonowaniu, które pojawiły się po kryzysie lat 2008-2011, jest zasadny. Sformułowanie hipotez badawczych także jest prawidłowe. Co należy podkreślić, wywód prowadzony w pracy jest klarowny, spójny i stanowi konsekwentną realizację założonego celu, część empiryczna zaś zawiera przeprowadzoną zgodnie ze sztuką weryfikację wspomnianych hipotez.

2. Cel i hipotezy badawcze pracy

We wstępie autor stwierdza, iż celem pracy jest porównanie zasadniczo dwóch podejść do pomiaru ryzyka portfolio kredytowego, a dokładnie podejścia zawartego w różnych „klasycznych” metodach używanych we współczesnej bankowości z podejściem prezentowanym w ramach metod wykorzystujących sztuczną inteligencję [s.9]. Jak wspomniałem, takie sformułowanie celu należy uznać za prawidłowe, przy czym, co istotne, jego realizacja wiąże się z – finalnie udanym - przedstawieniem nowatorskiego podejścia do tychże metod, którym jest właśnie szacowanie ryzyka portfolio kredytowego przy użyciu



technik uczenia maszynowego w postaci maszyny wektorów wspierających (SVM) i regresji wektorów wspierających (SVR). Inaczej powyższe kwestie ujmując, autor postawił sobie za zadanie dokonanie porównania zaproponowanego przezeń modelu z modelami aktualnie wykorzystywanymi w bankowości, przy czym kryterium oceny w tymże porównaniu zastosowanym miał być sukces predykcyjny względem szacowania ryzyka kilku wybranych portfolio kredytowych [„The aim of the thesis is to introduce a credit portfolio model which is effective (compared to a linear model) and even a better predictor than current credit portfolio models for some common bond portfolio”, s. 12].

W tej perspektywie sformułowanie hipotez badawczych wydaje się logiczną konsekwencją tak ujętego celu. Autor zaproponował trzy takie hipotezy, pierwszą, wedle której regresja wektorów wspierających stanowi efektywny model dla pomiaru ryzyka kredytowego typowych portfolio, drugą, głoszącą iż SVR wykazuje się mniejszym poziomem błędu wyrażonym w kategoriach RMSE niż wieloczynnikowy model liniowy (LM), i wreszcie trzecią, w myśl której SVR przewyższa skutecznością predykcyjną w szacowaniu ryzyka dla typowych portfolio również modele przemysłowe CreditRisk+® i, w szczególności, CreditMetrics®.

Ogólnie, w odniesieniu do celów i hipotez badawczych, należy stwierdzić, że zarówno ich wybór, jak też sformułowanie świadczy o daleko idącej dojrzałości badawczej autora. Ma on bowiem świadomość, iż proponuje nowe podejście do bardzo istotnych dla funkcjonowania instytucji finansowych kwestii, potrafi postawiony problem przełożyć na „język” badań empirycznych, przeprowadzić te ostatnie w sposób technicznie kompetentny i, co równie ważne, uogólnić wyprowadzone z tychże badań wnioski tak, by – w zakończeniu - zaproponować nowe kierunki badań. Pod tym względem oceniam więc zarówno kompetencje autora, jak i sposób realizacji postawionego sobie przez niego zadania, niezwykle wysoko.

### 3. Struktura pracy

Praca składa się ze wstępu, pięciu rozdziałów, zakończenia, zawierającej 1202 pozycje bibliografii oraz trzech aneksów. Bardzo przydatny w lekturze jest również załączony wykaz stosowanych w pracy skrótów.

Układ pracy jest klarowny, logiczny i w sposób widoczny podporządkowany realizacji założonego celu badawczego. W tym miejscu chciałbym podkreślić coś, co być może bezpośrednio nie wiąże się z – podlegającą ocenie – wartością naukową pracy, niemniej stanowi dodatkową jej zaletę. Otóż, swego rodzaju „pomysł” na układ tekstu, jak też dbałość autora o precyzyjne zdefiniowanie używanych w tymże tekście pojęć i terminów sprawia, iż omawiana praca ma także walor dydaktyczny i może służyć, zwłaszcza jej część teoretyczna, jako kompleksowe wprowadzenie do dziedziny badawczej, w ramach której realizowane są zawarte w pracy badania. Jest to okoliczność nie bez znaczenia, zwłaszcza jeśli praca ta miałaby zostać opublikowana, co w tym miejscu, również z uwagi na powyższą jej zaletę, gorąco rekomenduję.

W odniesieniu zaś do samej struktury autor konsekwentnie zaczyna w niej od (w rozdziale pierwszym) zdefiniowania podstawowych pojęć m. in.: aktywów (s. 17), inwestycji (s. 18), instrumentów kapitałowych (s. 19), obligacji (s. 20), portfolio inwestycyjnych (s. 34) (każdorazowo przedstawiając klasyfikacje wyżej wymienionych), by następnie omówić obowiązujące obecnie regulacje, którym podlega funkcjonowanie instytucji finansowych; w dalszej części rozdziału przedstawiona została klasyfikacja alternatywnych aktywów i funduszy inwestycyjnych, strategii inwestycyjnych, a także zdefiniowane zostały podstawowe, obecne w

literaturze przedmiotu, miary służące do pomiaru efektywności inwestycji i ich ryzyka (RARORAC, VAR).

Rozdział drugi zawiera omówienie terminologii związanej z zarządzaniem ryzykiem oraz prezentację kluczowych teorii portfelowych (Markovitz, CAPM, ATP) i z tego względu autor rozpoczyna tenże rozdział od zdefiniowania samego ryzyka i ogólnej idei jego pomiaru (s. 65), by następnie podkreślić szczególną - dla prowadzonych w pracy badań - rolę ryzyka kredytowego (s. 66). W dalszej części przedstawione zostają inne - niż to kredytowe właśnie - rodzaje ryzyka (m. in. rynkowe, stóp procentowych, operacyjne, walutowe, płynnościowe) oraz zdefiniowane zostaje zarządzanie ryzykiem (s. 80); autor omawia również przebieg samego procesu tego zarządzania (s. 81-82). Kolejnym zagadnieniem poruszonym w tym rozdziale są, szczegółowo omówione, miary ryzyka, co pozwala - w końcowej jego części - na omówienie modeli: Markovitz, CAPM i ATP.

W rozdziale trzecim scharakteryzowane zostały systemy ratingowe, w szczególności stosowane w tychże systemach kryteria (ilościowe i jakościowe), przebieg procedur ratingowych oraz ich rozróżnienie na zewnętrzne i wewnętrzne. Co istotne, szczególnie w perspektywie roli i znaczenia ratingów w ocenie ryzyka portfelowego, autor omawia również funkcje przez te ratingi pełnione oraz związane z ich stosowaniem zagrożenia takie, jak choćby możliwość wzmocnienia poprzez ich publikowanie cykliczności gospodarki (s. 145).

Rozdział czwarty zawiera prezentację „klasycznych” modeli portfelowych (CPM): modeli stopy ryzyka (których komercyjnym reprezentantem jest CreditRisk+®), modeli strukturalnych, inaczej Mertonowskich (z CreditMetrics® i KMV® na czele) oraz ekonometrycznych (CreditPortfolioView®), a także porównanie ich głównych zalet, wad, a także głównych obszarów zastosowań.

Kluczowy dla realizacji celu badawczego i weryfikacji hipotez badawczych jest natomiast rozdział piąty. Autor rozpoczyna go od przybliżenia samej koncepcji sztucznej inteligencji, jej „słabej” i „mocnej” wersji, uczenia maszynowego (i jego typów) jako jej szczególnej formy oraz pokrótce wyjaśnia, skąd mogą wynikać przewagi w zastosowaniu AI do analizy ryzyka portfolio kredytowego. W dalszej części szczegółowo omawia także konstrukcję i ideę stosowania SVM i SVR. Podrozdziały 5.3 (s. 245) i 5.4 (s. 257) już bezpośrednio referują przebieg i rezultaty badań empirycznych, których efektem jest weryfikacja hipotez badawczych. Po wyjaśnieniu struktury i reguł doboru czterech portfolio, które objęte były badaniem oraz skonstruowaniu zbiorczej skali ratingowej autor przedstawia proces weryfikacji pierwszej z hipotez zerowych: „ $H_0$ : The predicted distributions of values from LM and SVR are equal (stemming from the same population)” (s. 255). Wskutek wykazania za pomocą testu Kruskala-Wallisa istotnych rozbieżności w przewidywanych rozkładach dla wszystkich czterech portfolio hipoteza ta zostaje odrzucona. Podobnie rzecz ma się z drugą z hipotez zerowych: „ $H_0$ :  $RMSE(LM, \text{real portfolio data}) \leq RMSE(SVR, \text{real portfolio data})$  (s. 256), gdyż okazuje się, iż „błąd” jest istotnie mniejszy w przypadku SVR. Odrzucenie dwóch pierwszych hipotez jednoznacznie wskazuje, co podkreśla autor, iż SVR wykazuje znacząco większą, względem modelu liniowego, przydatność i efektywność w pomiarze ryzyka portfolio kredytowego (s. 257: „Hence, the support vector machine AI – model is superior to the linear model and principally useful and effective to measure credit portfolio risk”). Trzecia, z punktu widzenia założonego w pracy celu i w świetle tejże pracy tytułu, najważniejsza hipoteza zerowa, głosząca iż „ $H_0$ :  $\text{Min}(\text{CreditMetrics}^\circledast \text{ VaR} - \text{real VaR}; \text{CreditRisk}^\circledast \text{ VaR} - \text{real VaR}) \leq (\text{SVR VaR}, \text{real VaR})$  również, dla większości portfolio zostaje odrzucona, przy czym

*Gdańsk*

w przypadku CreditRisk+® SVR wykazuje wyższość, w kategoriach zbieżności przewidywanego VaR z VaR rzeczywistym, dla wszystkich czterech portfolio, a w przypadku CreditMetrics® wyższość ta ma miejsce w dwóch przypadkach, a w jednym zarówno CreditMetrics® jak też SVR w równym stopniu prognozują VaR niższy od realnego.

W zakończeniu autor nie tylko krótko podsumowuje przeprowadzone w pracy badania, ale też formułuje ciekawe sugestie względem ewentualnych dalszych kierunków badań nad wykorzystaniem AI do szacowania ryzyka kredytowego.

Podsumowując tę część oceny pracy pragnę ponownie podkreślić, że omówiona powyżej struktura, zawartość merytoryczna poszczególnych rozdziałów, jak też sam sposób prowadzenia wywodu, zasługują na najwyższe uznanie. Konsekwencja i spójność w tej strukturze dostrzegalne świadczą o tym, iż autor w sposób niezwykle świadomy i przemyślany realizuje założone w pracy cele. Na uwagę zasługuje także elegancja językowa, gdyż praca napisana jest w sposób, jak na tak złożoną jej materię, niezwykle zrozumiały.

#### 4. Najważniejsze osiągnięcia pracy

Za główne dokonanie autora uznać należy część empiryczną recenzowanej dysertacji. Począwszy od sformułowania hipotez badawczych, poprzez dobór podlegających badaniu portfolio, na sposobie weryfikacji hipotez i umiejętności oceny wynikających z tej weryfikacji wniosków kończąc, autor wykazuje się dojrzałością badawczą, która dalece wykracza poza wymogi stawiane tego typu pracom rygorowym.

Co istotne, i czego autor zdaje się mieć pełną świadomość, przeprowadzone badania – choć w swej istocie nowatorskie – stanowią właściwie wstęp do dalszych analiz, które mogą skutkować istotnym przewartościowaniem stosowanych w obszarze szacowania ryzyka kredytowego metod. W tym sensie docenić należy także sformułowane w zakończeniu uwagi, które wskazują na zalety metod opartych na sztucznej inteligencji oraz na podstawy ich przewagi nad metodami „klasycznymi”.

Wartość części empirycznej nie umniejsza jednak tego, iż na uznanie zasługuje także część teoretyczna pracy, gdyż, jak już wspomniałem, może stanowić ona, dydaktycznie niezwykle wartościowe, wprowadzenie do omawianych zagadnień.

#### 5. Uwagi krytyczne i polemiczne

W tej części niniejszej recenzji pozwolę sobie zgłosić kilka uwag krytycznych względem pracy oraz poczynić kilka refleksji, które – głównie jako filozofowi ekonomii – jawią mi się jako warte zasygnalizowania. Jak łatwo wywnioskować z wcześniej formułowanych ocen, ani te uwagi krytyczne, ani tym bardziej kwestie polemiczne, nie zmieniają mojej niezwykle wysokiej oceny całej pracy i w tym sensie należy je traktować raczej jako asumpt do dyskusji, bądź też sugestie względem dalszych badań, które autor z pewnością będzie prowadził.

Przede wszystkim zatem, zwłaszcza w perspektywie publikacji omawianej pracy (co, jak już zaznaczyłem, uważam za wskazane), stanowczo za krótki wydaje się rozdział piąty. Badania empiryczne stanowią bowiem najistotniejsze dokonanie autora i o ile nie można ich omówieniu zarzucić braku klarowności, to z pewnością jest ono w pewnym sensie zbyt pobieżne i poniekąd można by błędnie z tego wywnioskować, iż stanowi ono jedynie dodatek do części teoretycznej pracy (gdy tymczasem jest wprost przeciwnie).



W odniesieniu zaś do zagadnień bardziej szczegółowych muszę zwrócić uwagę, iż twierdzenie, wedle którego „inflation, as commonly known in economics and, e. g. shown by Friedman is mainly a monetary phenomenon” (s. 44) nie odpowiada poglądom ogółu ekonomistów na tę kwestię. Już choćby w obecnie prowadzonych w tym obszarze sporach da się zauważyć - wśród wymienianych przez ekonomistów czynników powodujących inflację - silną obecność czynników innych niż pieniężne (kosztowych, popytowych, wynikających z wpływu polityki na oczekiwania inflacyjne, itp.)<sup>1</sup>. Mając świadomość, iż nie stanowi to głównego wątku rozważań, uznaję jednak za istotne by autor tę kwestię uwzględnił w swych dalszych badaniach, zwłaszcza, że zdaje się to czynić kwestię szacowania związanego z tym ryzyka dużo bardziej złożoną.

Drugą - z mojej perspektywy ciekawą i mającą już bardziej polemiczny charakter – kwestią jest, mające w swej istocie źródło w postulowanej ontologii, rozróżnienie zawarte w następującym sformułowaniu: „the security margin hence is the difference of the (traded) share price to the share’s inner value or ‘fair’ price” (s. 58). Otóż, byłbym ciekaw rozwinięcia tego rozróżnienia, a w szczególności przybliżenia ewentualnego znaczenia owych „wewnętrznej wartości” i „uczciwej (sprawiedliwej?) ceny”. Kwestia ta wydaje się o tyle istotna, że jeśli takowe rozróżnienie dałoby się przeprowadzić, (wobec czego - jako radykalny konstruktywista - jestem dalece sceptyczny), to miałyby to znaczący wpływ na sposób wykorzystania metod opartych na sztucznej inteligencji.

Podobne w swym charakterze, choć tu bardziej w swej istocie epistemologiczne, zagadnienie stanowi rozróżnienie obecne w podrozdziale 2.2, który dotyczy miar ryzyka. Autor stwierdza w nim, iż „in order to measure risk, one needs to define, in a subjective or objective way, the range of possible outcomes (events) and assigned probabilities” (s. 87). Być może moja wątpliwość wynika z niedostatecznej wiedzy w tym zakresie, ale tym bardziej intryguje mnie kwestia tego, czym różnią się subiektywny i obiektywny sposób definiowania zakresu możliwych wyników?

Kolejne dwa zagadnienia, które w mojej opinii warte byłyby doprecyzowania lub rozwinięcia, związane są z uwagami sformułowanymi przez autora w odniesieniu do ratingów. Pierwsze z nich dotyczy stwierdzenia, iż „such a rating ought to be constructed in a way that the following axioms hold truth [...]. The rating system is (fairly) objective [...]” (s. 125). Intryguje mnie w tym kontekście, jak owe „prawda” i „obiektywność” ratingów są rozumiane i w konsekwencji, jak oceniane jest to czy dany rating je osiąga. Drugim zaś problemem, który moim zdaniem zasługuje na uwagę, jest zasygnalizowana przez autora swoista procykliczność ratingów, a dokładnie to, iż mogą one – zwłaszcza w okresach recesji – wzmacniać negatywne tendencje gospodarcze (s. 145). Oczekiwania, ich kształtowanie, wpływ wiedzy o poziomie różnego rodzaju wskaźników (czy jak w tym przypadku znajomość ocen ratingowych) oraz swego rodzaju sprzężenia, które w tym względzie zachodzą, to kwestie bardzo aktualne nie tylko we współczesnej ekonomii, ale w ogóle w naukach społecznych. Pytanie zatem jest następujące: czy, mając świadomość wpływu ratingów na działania podmiotów, można by traktować je poniekąd jako narzędzia zarządzania ryzykiem?

Następną kwestią, która w mojej opinii warta jest poruszenia, to rozróżnienie ryzyk: systemowego, systematycznego i idiosynkratycznego (s. 188). W szczególności chodzi mi o

---

<sup>1</sup> Choćby: Shapiro A. H., *How Much Do Supply and Demand Drive Inflation?*, FRBSF Economic Letter 2022-15, July 2022.

zestawienie tego rozróżnienia z obecnym – za sprawą F. Knighta - w teoretycznym dyskursie w ekonomii podziałem na ryzyko (które jest mierzalne) i, w swej naturze niemierzalną, niepewność<sup>2</sup>. Pytanie jest następujące: czy można by utożsamić ryzyko systemowe, zdefiniowane jako „the risk which lies in the entire financial system” (s. 188), z ową Knightowską, niemierzalną niepewnością?

Wreszcie chciałbym też zasygnalizować problem trochę bardziej złożony, który jest w swej naturze problemem metodologicznym, a który o tyle zdaje się być istotny, że zastosowanie sztucznej inteligencji może tu wiązać się ze zmianą sposobu rozumienia nauki jako pewnej dziedziny działalności ludzkiej. Otóż, o ile rozumiem uwagi zawarte w zakończeniu (s. 269), zastosowanie maszyny wektorów wspierających ma tę zaletę, iż „they ‘learn from patterns in data [...]’. As neither underlying a priori distribution assumptions nor potentially non-realistic constraints regarding asset price movements nor common parametric defaults intensities are necessary for SVR, not even an economically feasible model framework” (s. 269). Innymi słowy, modele oparte na, ogólnie to ujmując, AI charakteryzują się tym, że są niejako w swej naturze ateoretyczne. Konsekwencją tego jest, iż (co wyszczególnione zostało w tabeli na stronie 223) metody te (jako metody uczenia nadzorowanego) charakteryzują się przede wszystkim zdolnościami predykcyjnymi, przy ograniczonych zdolnościach wyjaśniania („prediction performance, and explainability is of limited importance”), a proces generowania danych pozostaje „nieznany”. Interesuje mnie czy autor zgodziłby się ze stwierdzeniem, iż zastosowanie metod opartych na AI mogłoby doprowadzić do swego rodzaju rozstrzygnięcia „sporu o metodę” na rzecz instrumentalizmu, czyli poglądu, wedle którego nieistotne są założenia modelu/teorii i w konsekwencji wyjaśnienia sposobu funkcjonowania analizowanych mechanizmów/obszarów działalności ludzkiej, a wartość naukowa wszelkich konstrukcji sprowadza się do ich zdolności predykcyjnych jedynie.

Na zakończenie tej części recenzji pragnę jedynie powtórzyć, iż wszelkie uwagi w niej sformułowane w żaden sposób nie wpływają na pozytywną ocenę pracy i wynikają raczej z tejże pracy heurystycznej płodności niż jakkolwiek rozumianych jej słabości.

#### 6. Uwagi końcowe. Konkluzja.

Recenzowana dysertacja stanowi niewątpliwie znaczące osiągnięcie naukowe. Autor wykazał się w niej zarówno umiejętnością analizy zagadnień teoretycznych, jak też daleko idącą dojrzałością w obszarze badań empirycznych. Nie tylko zaproponował całkowicie nowe podejście do niezwykle istotnych w obszarze finansów kwestii, ale też potrafił empirycznie dowieść jego przewagi nad rozwiązaniami powszechnie w tymże obszarze stosowanymi. Z tego też względu, jednocześnie sugerując jej opublikowanie, stwierdzam, iż zgodnie z art. 187 ustawy z dnia 20 lipca 2018 *Prawo o szkolnictwie wyższym i nauce* (Dz. U. 2021 poz. 478) recenzowana praca spełnia wszelkie wymogi stawiane w w/w ustawie i na tej podstawie wnioskuję o jej przyjęcie i dopuszczenie do publicznej obrony. Wnoszę także o jej wyróżnienie.

---

<sup>2</sup> Knight F., *Risk, Uncertainty and Profit*, Beard Books, Washington DC, 2002.