

## Posudek oponenta diplomové práce

### Vojtěch Herrmann: Moderní predikční metody pro finanční časové řady

Předložená diplomová práce se zabývá modelováním a predikcí časových řad pomocí gradientně boostovaných rozhodovacích stromů (XGBoost podle příslušné softwarové knihovny). Jedná se o moderní přístup zejména v rámci finančních časových řad (na rozdíl od tradičních modelů typu ARIMAX a GARCH), který souvisí s oblastí strojového učení. Kromě porovnání obou zmíněných přístupů je v práci popsána také jejich kombinace do podoby hybridního modelu.

Uspořádání práce odpovídá logicky jejímu obsahu. První kapitola popisuje teoretický rámec tzv. supervizovaného učení a modelů ARIMAX. V druhé kapitole je představen datový soubor pro aplikaci dané metodiky (jedná se o časovou řadu denních zobchodovaných objemů indexu S&P 500, který je velmi frekventovaný v praxi). Třetí až pátá kapitola se zabývá predikcemi pomocí ARIMAX a XGBoost modelů. Konečně šestá kapitola je věnována kombinovaným hybridním modelům.

Co se týče formální podoby, diplomová práce má nadstandardní rozsah blízký rozsahu disertací (téměř sto stran). To je částečně dáno tím, že práce je silně počítačově orientovaná a vyžaduje prezentaci mnoha počítačových kódů a grafiky. Je otázkou, zda takové partie neměly být zařazeny do příloh diplomové práce (to by dle mého názoru přispělo k její lepší přehlednosti, aniž by došlo k významné redukci sdělovaných informací). Vystává zde také často problém terminologie, např. s překladem anglického termínu do češtiny nebo s nekonzistencí terminologických zdrojů v literatuře. Diplomant používá velké množství anglikanismů, které by byly neúnosné i v pracovní komunikaci mezi IT pracovníky, natož pak v akademické práci (např. „čím přesnější fit, tím...“, „objective je tedy...“, „maximalizace gainu“, „trénování rezultující v aditivní predikci“, „naši představu potvrdíme při grid searchi“), někdy ponechává v textu anglický termín, jindy ho opíše českým výrazem. Navíc se používají některé nestandardní jazykové obraty (i v názvech odstavců, kde zní velice nepřirozeně, např. 1.3. Teorie vedoucí k XGBoostu, str. 16: Přesný hladový algoritmus). Při tomto stylu vyjadřování měl diplomant raději práci napsat v angličtině, neboť ve stávající podobě se čte velice obtížně.

Po obsahové stránce by dle mého názoru mohla být práce obhajovaná i před informatickou komisí. O tom svědčí mimo jiné i referenční zdroje využívané v práci, což jsou často sborníky z převážně informatických konferencí publikované on-line (např. *International Conference on Information and Communication Technologies*). Diplomová práce často přebírá styl takových příspěvků a stává se méně srozumitelnou pro neodborníka v této oblasti. Na druhé straně je nutné ocenit, že diplomant byl schopen se zorientovat v takto prezentované problematice a byl schopen extrahovat její statistické či optimalizační aspekty. V každém případě je tato diplomová práce důkazem, že její autor vykazuje v čistě softwarových záležitostech solidní dovednosti.

K práci mám následující komentář a dotazy:

4<sup>11</sup>: Jednokrokové předpovědi postačují jen pro investory se silně krátkodobým předpovědním horizontem. Zkoušel diplomant aplikovat danou předpovědní metodiku také pro vícekové předpovědní horizonty, kam vstupuje vliv trendů, sezónnosti apod.?

22<sub>10-1</sub>: Velkým nedostatkem dané metodiky je dle mého názoru absence predikčních intervalů pomocí XGBoost, protože z praktického pohledu jsou tyto intervaly v daném kontextu

velmi důležité. Diplomant se nepokusil v rámci provedené numerické studie tuto mezeru doplnit (např. pomocí zmiňovaného neparametrického bootstrapu)?

31<sup>9,8</sup>: Není v rámci uváděné numerické studie vhodnější alternativou GARCH model s exogenními proměnnými než použitý model ARIMAX (viz zmiňovanou heteroskedasticitu)?

55<sup>1-3</sup>: Jak lze interpretovat výsledky XGBoostu pomocí Shapleyho hodnot (jedná se patrně o nástroj používaný v kooperativních her)? Interpretace výsledků je totiž dost důležitá.

– Je daná metodika použitelná pro český akciový trh (při negativním stanovisku lze uvést hlavní příčiny, které tomu brání)?

Závěr:

Předložená práce dle mého názoru splňuje požadavky kladené na diplomové práce. Proto doporučuji, aby jako diplomová práce byla uznána.

19. 1. 2021

Prof. RNDr. Tomáš Cipra, DrSc.