

Posudek oponenta diplomové práce

Autor/Autorka

Bc. Veronika Báčová

Název práce

Oceňování v modelech stochastické volatility založené na hlubokém učení (Deep learning-based pricing in stochastic volatility models)

Studijní obor

N0541A170005 Matematika a finanční studia

Oponent práce

RNDr. Blanka Šedivá, Ph.D.

Splnění cílů práce:

nadstandardně velmi dobře splněny s výhradami nebyly splněny

Odborný přínos práce:

nové výsledky netradiční postupy zpracování výsledků z různých zdrojů shrnutí výsledků z různých zdrojů bez přínosu

Matematická (odborná) úroveň:

vynikající velmi dobrá průměrná podprůměrná nevyhovující

Věcné chyby:

téměř žádné vzhledem k rozsahu přiměřený počet méně podstatné, větší množství podstatnější, větší množství závažné

Grafická, jazyková a formální úroveň:

vynikající velmi dobrá průměrná podprůměrná nevyhovující

Slovní hodnocení a dotazy:

Cílem předkládané diplomové práce bylo popsat přístupy k oceňování opcí založené na hlubokých neuronových sítích. Aktuálnost uvedené problematiky dokládá například přehled využití neuronových sítí pro odhady cen opcí uvedených v článku [13] autorů J. Ruf a W. Wang z roku 2020, ve kterém jsou shrnuty nejvýznamnější publikace v této oblasti od roku 1993. Druhým cílem práce bylo implementovat ve vhodném prostředí (MATLAB) vybraný model a otestovat jej na reálných tržních datech. Deklarované cíle práce byly splněny velmi dobře.

Celkově považuji práci za velmi zdařilou, autorka kombinuje netradiční postupy a kombinace přístupů z různých zdrojů. Kromě simulační části, kdy pracuje se synteticky generovanými daty se velmi dobře vyrovnala i se zpracováním reálných dat. Práce je velmi zdařilá i z hlediska grafické zpracování, kdy autorka využila nově zaváděného celofakultního grafického vzoru na kvalifikační práci.

Vzhledem k rozsáhlosti studované problematiky bylo třeba soustředit se pouze na konkrétní partie, práce tak pracuje s Hestonovým modelem a jako reálná tržní data byl zvolen index SPX za období jednoho roku (2019). Z pohledu koncepce práce považuji za nadbytečné zařazení příkladu 1 (kapitol 2.4.4, 2.4.5 a kapitola 3.1). Ostatní kapitoly na sebe logicky navazují a tvoří ucelený text.

Kapitola 2 obsahuje základní definice a charakteristiky použitých metod. V této kapitole by bylo vhodné doplnit všechny dále používané aktivační funkce, v kapitole 2 (strana 15-17) jsou popsány a vykresleny tři vybrané, v implementační části práce (strana 29-30) je využito 5 základních typů aktivačních funkcí. V kapitole 3 je popsána metodologie přístupu k datům, za důležitější považuji subkapitoly věnované syntetickým datům dle Hestonova

modelu. Vhodné by bylo doplnit Tab. 3.3 o všechny parametry, které byly pro trénovací proces využity. Stejně tak by bylo vhodné doplnit kapitolu o histogram chyb (analogicky jako v kapitole 3.1, Obrázek 3.2).

Za nejdůležitější část práce považuji kapitolu 4, která obsahuje výsledky dříve popsanych metod a postupů aplikované na syntetická data a reálná data. Zejména část aplikace na reálná data ukazuje problematiku práce s reálnými daty, kdy je třeba věnovat zvýšenou pozornost přípravě dat i následné interpretaci. Autorka práce se s uvedenými problémy vypořádala výborně. Eliminace pozorování z datumu 18.12. 2019 však působí příliš jednorázově a účelově. Zajímavým doplněním přístupu založeného na neuronových sítích je porovnání výsledků získaných neuronovými sítěmi a klasickým přístupem, tj. odhadem parametrů modelu na základě nelineárních regresních modelů popsané v kapitole 4.2.3.

Grafická úroveň práce je velmi dobrá, práce neobsahuje téměř žádné překlepy a jiné formální nedostatky. Jako nevhodné však vidím umístování obrázků, v několika případech se stalo, že obrázky z jedné podkapitoly jsou zobrazeny až v následující podkapitole, což komplikuje orientaci v textu (např. Obrázek 2.5 příslušející k podkapitole 2.4.1 je zařazen až v kapitole 2.4.2). Elektronická verze práce obsahuje též zdrojová data a kódy. Ucelená informace uvedená v souboru Readme.txt o programech a datech, včetně informace o zdrojových souborech pro obrázky je velmi podnětná. Popis však v některých částech neodpovídá struktuře (adresář Data – calibration). U jednotlivých m-souborů také není zřejmé autorství, protože většina zdrojových kódů neobsahuje komentáře ani jiné informace o autorství. V některých případech jsou uvedeny soubory, které pravděpodobně „počítají“ stejné věci, např. myblsprice.m a BSeuCall.m. Orientaci v doložených souborech také znesnadňuje skutečnost, že v různých adresářích jsou soubory stejného názvu, není tak zřejmé, zda i obsahově jsou soubory stejného názvu shodné.

Při obhajobě navrhuji zaměřit se na následující otázky:

- Jakou výhodu má využití transformovaných parametrů, jako vstupní hodnot pro trénovací proces? Tj. proč bylo přistoupeno k úpravě vstupních dat dle kapitoly 3.2.2?
- Jak je zmíněno v práci, odhadnuté parametry pro index 241 (odpovídající datumu 18.12.2019) vykazují zjevné problémy. „Podezřelá“ je například i zápornost parametru volatility. U kolika dalších indexů došlo k obdobné situaci?
- Došlo k „problematické“ situaci s hodnotami parametrů modelu k datu 18.12.2019 i při odhadech pomocí nelineární regrese?

Předloženou práci doporučuji uznat jako kvalifikační diplomovou práci a navrhuji hodnocení známkou VÝBORNĚ.

Datum, jméno a podpis:

B. G. Lohr