

MATEMATIČKI FAKULTET
BEOGRAD

SEMINARSKI RAD

Tema:

***PRIMENA LSTM REKURENTNIH
NEURONSKIH MREŽA
ZA PREDVIĐANJE CENA NA BERZI***

Filip Šašić
Broj indeksa: 135/2015

1. Apstrakt

Rekurentna neuronska mreža (RNN) je model koji je pokazao izvrsne rezultate kod vremenskih serija i sekvencijalnih podataka u mašinskom učenju, gde je potrebno učenje distantnih zavisnosti medju ulaznim podacima. Primenuje se za prepoznavanje govora, generisanje teksta kao i u mnogim drugim oblastima. Ovaj rad istražuje način funkcionisanja rekurentnih mreža, konkretno long-short tern memory (LSTM) modela, koji donekle rešava problem nestajućeg i eksplodirajućeg gradijenta. Celu implementaciju pokušaćemo da primenimo na predviđanje cena na berzi, analiziraćemo rezultate i razmotriti mane i prednosti.

2. Uvod, opis problema i postojeći načini za rešavanje

Informacione tehnologije doprinele su ogromnom rastu metodološkog spektra raspoloživog donositeljima odluka u području poslovanja. Osim metoda operacionih istraživanja i statističkih metoda, veštačka inteligencija zasnovana na teoriji učenja unapredila je mogućnost korišćenja prethodnog znanja (kroz ekspertne sisteme) i podataka (kroz neuronske mreže) u cilju donošenja učinkovitih odluka. Naglasak ovog rada je na rekurentnim neuronskim mrežama (RNN) i upotrebi njihovih algoritama u oblasti finansija.

Neuronske mreže mogu se upotrebljavati za predviđanja, klasifikaciju i asocijaciju u različitim problemskim područjima. One su alat, namenjen radu s velikim količinama promenljivih podataka i podataka sa smetnjama, koji uključuju skrivene nelinearne zavisnosti.

Nedostatak standardiziranih paradigmi koje će odrediti efikasnost pojedinog algoritma i arhitekture u pojedinom problemskom području i na pojedinim karakteristikama podataka naglašeno je kod mnogih autora.

Konkretno u oblasti finansija neuronske mreže su pronašle svoje primenu u online berzama, s obzirom na novije izazove tehnologije sadašnjice i kreiranja budućih tokova. Značajna istraživanja se vrše u predviđanju deonice, kako bi se osigurala zarada korištenjem različitih statističkih metoda, tehničkih analiza među podacima.

Veštačke neuronske mreže se značajno primenjuju u istraživanjima poslovnih okruženja iz razloga da se kreira mogućnost učenja i detektovanja relacija među ne linearnim varijablama. Zapravo, veštačka inteligencija omogućava detaljniju analizu veće količine podataka, posebno onih podataka koji imaju tendenciju menjanja. Sve to upućuje na primenu veštačkih neuronskih mreža. Poslovno okruženje današnjice, prolazi kroz transformaciju poslovanja poslednjih deset godina i veštačke neuronske mreže su se pokazale kao model koji se u velikom broju koristi u biznisu, finansijama, ekonomiji s ciljem predviđanja serija performansi merenja. Brzine promena koje se dešavaju, danas putem virtuelnih rešenja su promenile tokove tradicionalnih berzi, a stanje na berzi je od posebnog značaja za investitore.

Finansije i ulaganja su oblasti u kojima se, prema nekim istraživanjima, najčešće primenjuju neuronske mreže. Neki od problema u čijem su rešavanju primenjivane neuronske mreže su predviđanje bankrota, procena rizika kod hipotekarnih i drugih vidova zajmova, predviđanje na berzama (predviđanje cena akcija, obveznica, trgovanje robama i dr.), različite finansijske prognoze itd. Neke od najpoznatijih kompanija koje uspešno primenjuju neuronske mreže za rešavanje svojih finansijskih i investicionih problema su Chase Manhattan Bank i American Express.

Rezultati velikog broja istraživanja, ali i praktičkih primena neuronskih mreža u poslovanju, pokazali su prednost neuronskih mreža u odnosu na klasične metode koje ne uključuje primenu veštačke inteligencije.

Prema nekim ranijim istraživanjima, najčešća područja primene neuronskih mreža u poslednjih 10 godina su proizvodnja i finansije. Neuronske mreže u finansijama se najčešće upotrebljavaju za predviđanje rezultata ulaganja u deonice, kao i za izbor deonica u trgovanju na berzi. Iako ne postoje standardizovane paradigme koje bi odredile efikasnost pojedinog algoritma neuronske mreže i njene arhitekture u određenoj oblasti, prilikom primene neuronskih mreža na tržištima deonica primenjuje se samo jedan algoritam.

Ukoliko bi posmatrali samo analizu na tržištima deonica, možemo uvideti da postoje tri glavne grupe problema u kojima se primenjuju neuronske mreže.

- 1) Prva grupa se sastoji od predviđanja ponašanja deonica klasifikujući ih u klase kao što su: deonice sa pozitivnim i sa negativnim povraćajem ili klase deonica koje će imati dobar, neutralan ili loš rezultat. Takva primena neuronskih mreža pruža vrednu podršku u donošenju investicionih odluka, ali ne specificiraju tačnu vrednost očekivane cene ili očekivanog profita.
- 2) Druga grupa se sastoji od predviđanja cena deonica. Takvi sistemi nastoje da na temelju prethodnih cena deonica i povezanih finansijskih pokazitelja predvide cenu deonica jedan ili više dana unapred.
- 3) Treća grupa je orijentisana na modelovanje ponašanja deonica i na prognoze. Takve aplikacije nisu orijentisane samo na predviđanje budućih vrednosti, već i

na procenu značaja pojedinih faktora, analizu osetljivosti između varijabli koje utiču na sam rezultat, kao i na druge analize međusobnih zavisnosti.

I pored široko rasprostranjene upotrebe neuronskih mreža u finansijama i investiranju, još uvek postoje otvorena pitanje i smernice za dalja istraživanja.

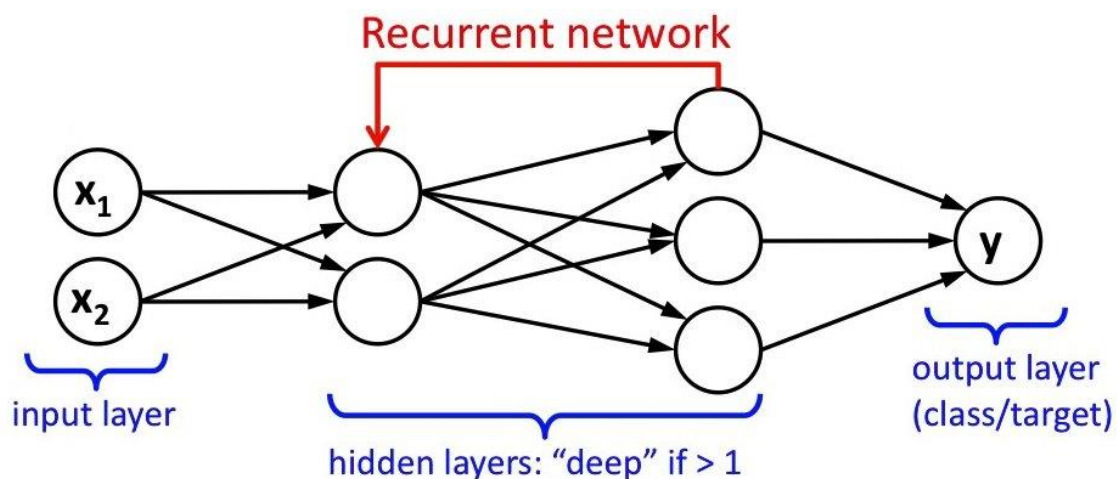
Dizajn modela podatak za neuronske mreže se određuje izborom ulaznih i izlaznih promenljivih. Iako se upotrebljavaju različiti modela podataka, nijedan se ne može smatrati predominantnim, pa ova raznolikost može uzrokovati teškoće u konstruisanju paradigmi o efikasnosti neuronskih mreža. Broj ulaznih promenljivih u neuronskim mrežama varira , od kojih su najčešće cena deonica i različiti finansijski koeficijenti.

Uvođenje kvalitativnih podataka je novi pristup aplikacijama neuronskih mreža i otvara nove mogućnosti za dalja istraživanja. Napredak je postignut 1997. godine kada su Sepp Hochreiter i Jurgen Schmidkuber predložili korsicnje novog modela nazvanog long-short term memory (LSTM) koji je rešavao dotada poznati problem nestajućeg i eksplodirajućeg gradijenta.

3.Rekurentne neurnoske mreže (RNN)

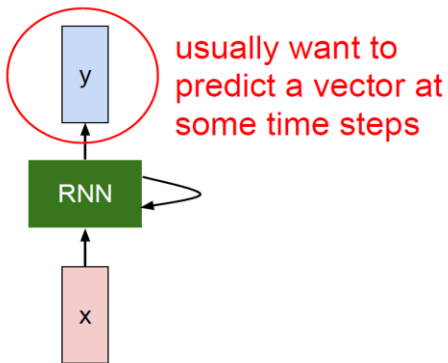
Hopfield-ova neuronska mreža pripada grupi rekurentnih neuronskih mreža. Kod ovih mreža, izlazni neuroni su povezani sa svim ulaznim neuronima, osim sopstvenog. Na taj način svaki od neurona svojim stanjem direktno utiče na sve ostale u mreži.

Elman-ove neuronske mreže su poznate i pod nazivom parcijalne rekurentne neuronske mreže ili proste rekurentne neuronske mreže. Reč je višeslojnim preceptronima proširenim sa jednim ili više kontekst slojeva koji pamte izlazne vrednosti jednog od slojeva koji kasni za jedan korak. Ovi slojevi se koriste za aktivaciju tog ili nekog drugog sloja u sledećem vremenskom koraku. U ovom radu baviceemo se ovom vrstom mreža.

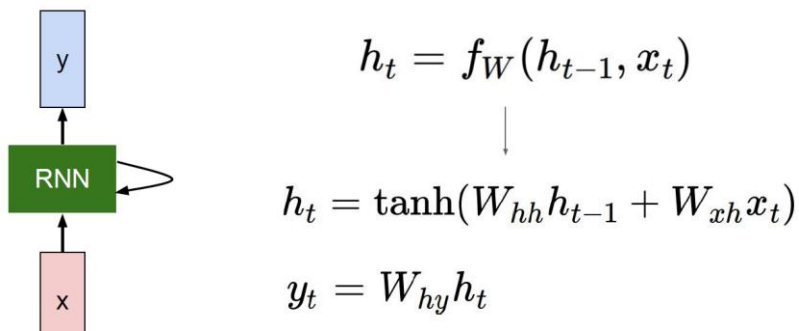


3.1 Arhitektura

Kod RNN obicno želimo da predvidimo vektor izlaznih vrenosti nakon nekih vremenskih koraka.

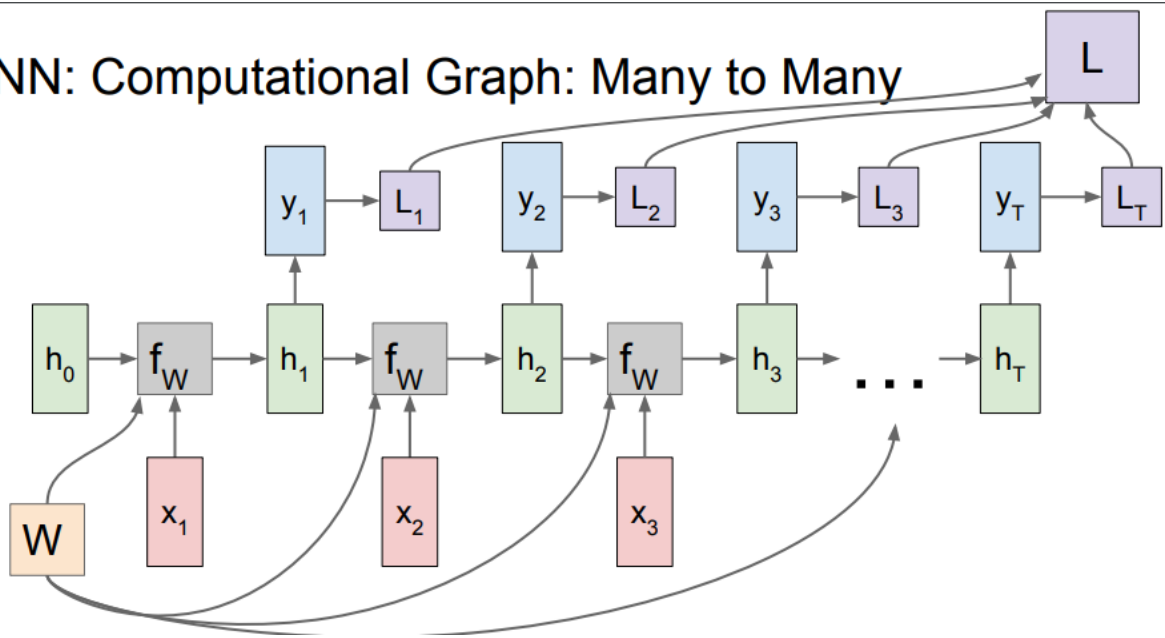


Skriveno stanje se sastoji od jednog vektoroa h . Ovo stanje zavisi od trenutnog ulaza X_t i predhodnog skrivenog stanja h_{t-1} .



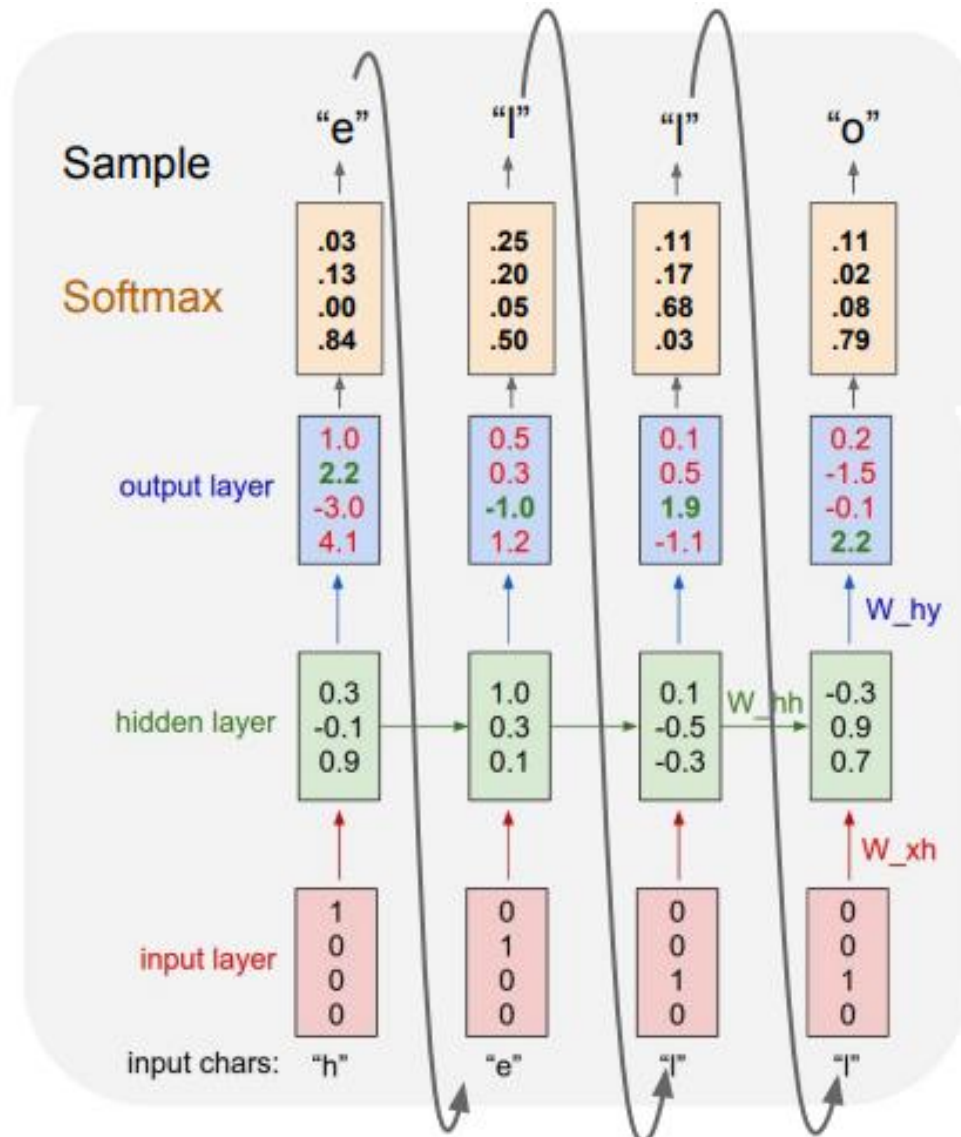
Rekurentnu neuronsku mrežu prvo razvijamo u određen broj koraka k , i u svakom koraku koristimo iste vrednosti za težinske matrice W . Za svaki korak računamo gresku L_k i sabiranjem ovih vrednosti dobijamo ukupnu gresku.

RNN: Computational Graph: Many to Many



Svaki korak prima nove ulazne vrednosti X_k i h_{k-1} . Koriste se iste težinske matrice W za svaki korak, a zatim se računaju izlazne vrednosti Y_k i greske L .

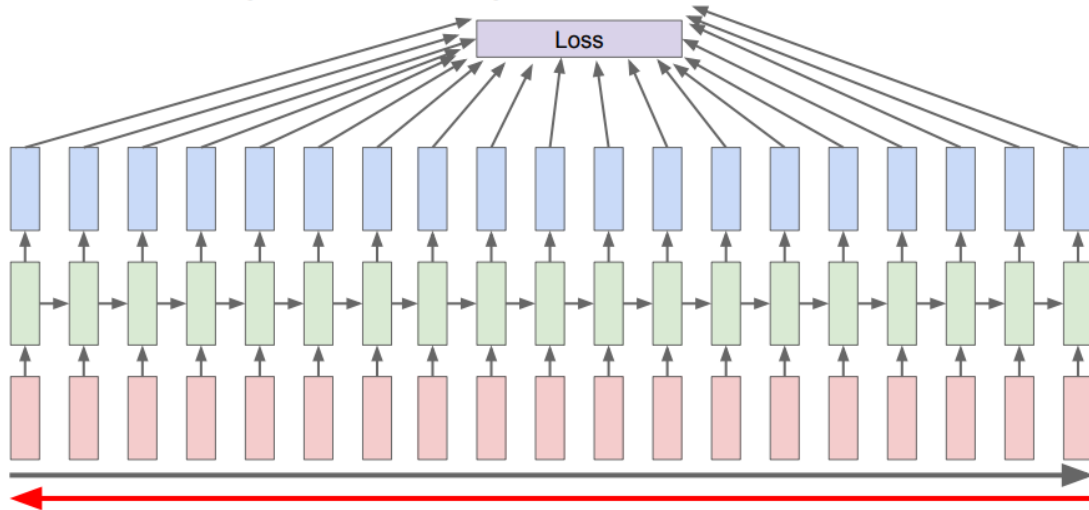
Na sledećoj ilustraciji vidimo konkretan primer razvoja RNN-a za predviđanje karaktera. Cilj ove mreže je da predividi karakter 'o' u reci 'hello'



Kada razvijemo mrežu do kraja, naredni korak je propagacija unazad kroz vreme (backpropagation through time). Počevši od poslednjeg koraka pa sve do prvog,

računamo negativni gradijent za svaki korak, zatim vrši se suma gradijenata i deli sa brojem koraka kako bi se dobio prosečan gradijent za tu iteraciju (razvoj mreže). Ovde se takodje vrše jos neke transformacije gradijenta kako bi se ovaj korak optimizovao konkretno za problem koji rešavamo.

Backpropagation through time



3.2 LSTM – long short term memory rekurentne mreže

Rekurentne mreže koje se sastoje od LSTM jedinica zovu se još i LSTM mreže. Nastale su 1997. godine kada su Sepp Hochreiter i Jürgen Schmidhuber predložili ovaj model kao izmenu za postojeći model standardnih RNN. Ovaj model donekle rešava problem nestajućeg i eksplodirajućeg gradijenta tako što omogućava drugaciji tok podataka kroz svoju ćeliju.

3.2.1 Arhitektura

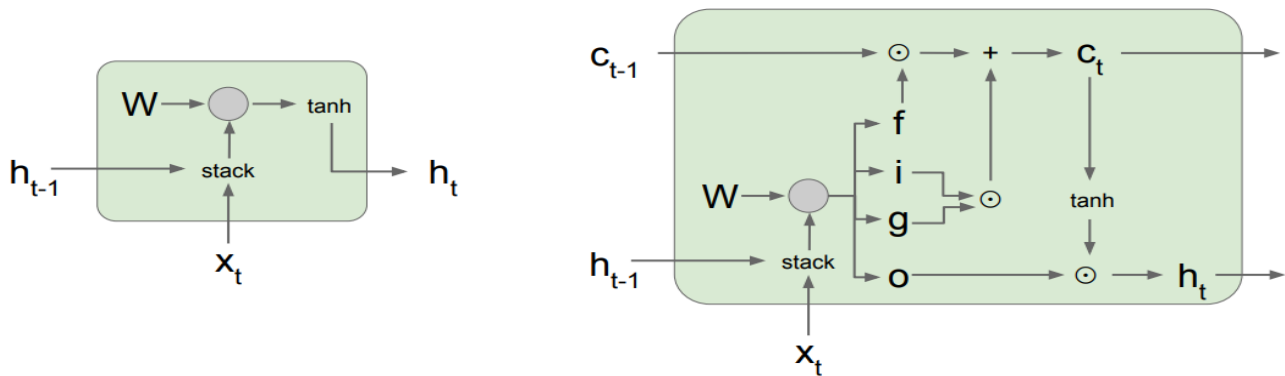
LSTM jedinica koja zamenjuje standardnu RNN jedinicu, sastoji se od četiri kapije(gates) koje kontrolišu prohodnost podataka(input gate, forget gate, output gate, gate gate).

- 1) Kapija zaboravljanja (forget gate) – nam govori da li treba da obrišemo ćeliju
- 2) Ulazna kapija (input gate) – nam govori da li da pišemo u ćeliju

- 3) Izlazna kapija (output gate) – nam govori u kojoj meri da otkrijemo ćeliju spoljašnosti
- 4) Kapija kapija (gate gate) – nam govori koliko da pišemo u ćeliju, ova kapija ima promenljivo ima u literaturi.

Svaka od ovih kapija je ustvari mala neuronska mreža za sebe veličine: $(x_t + h_{t-1}) \times h_t$
Ove male mreže imaju svoje težinske matrice i nemaju skrivene slojeve.

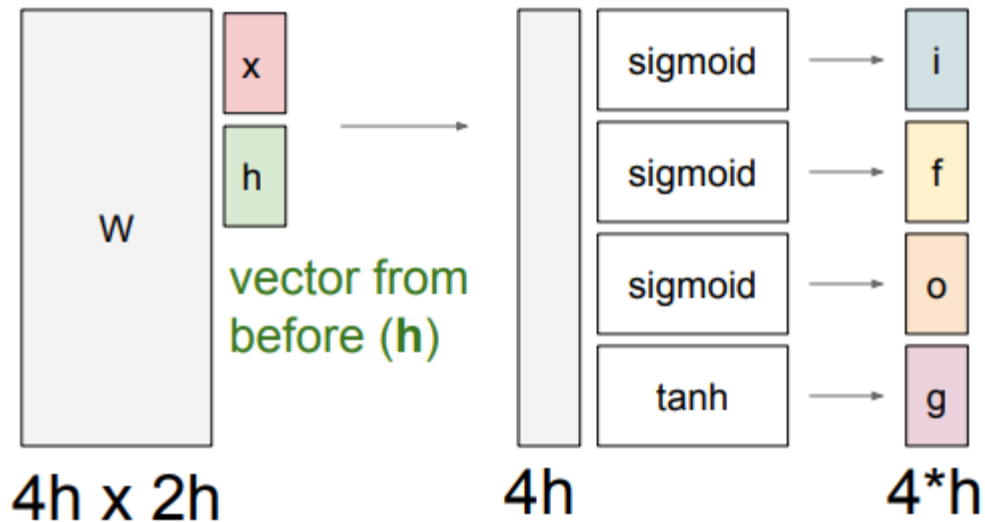
Pored ovih kapija LSTM jedinica sadrži dva stanja, ćelijsko stanje(cell state) i skriveno stanje(hidden state) koje ujedno predstavlja izlaz iz LSTM jedinice



Standardna RNN vs LSTM jedinica

Kao što se vidi sa dijagrama na stanje ćelije C_t ne utiču direktno algebarske operacije težinskih matrica i to u nekoj meri sprečava problem da gradijent bude previše mali ili previše veliki.

Treba napomenuti da operacije množenja i sabiranja sa dijagrama su operacije element po element (elementwise)



Radi lakšeg razumevanja, često se u literaturi četiri težinske matrice svake of kapija predstavljaju kao jedna velika W . Na ovoj ilustraciji predpostavlja se da je velicina ulaznog vektora x ista kao velicina vektora skrivenog sloja h

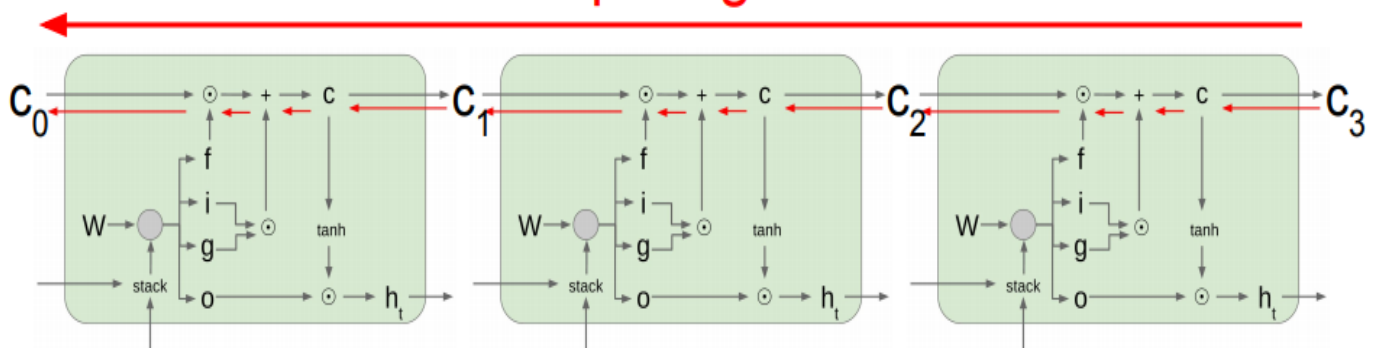
Formule za propagaciju unapred kroz LSTM mrežu glase:

$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}$$

$$c_t = f \odot c_{t-1} + i \odot g$$

$$h_t = o \odot \tanh(c_t)$$

Uninterrupted gradient flow!



4. Primena LSTM mreže na problem predviđanja cene na berzi i rezultati

U programskom jeziku Python implementirali smo LSTM mrežu korišćenjem samo bililoteke Numpy, bez dodatnih pomoćnih biblioteka i funkcija. Glavni cilj ovog projekta je da se nauče precizni koraci i tačna implementacija LSTM mreža, sa akcentom na propagaciju unapred i unazad kroz vreme.

Akcenat nije na rezultatima zbog prirode problema (berze su veoma teško predvidive, neki stručnjaci šaljivo tvrde da majmun koji baca strelicu u pikado i na osnovu toga bira da li će da kupi ili proda berzu ima veće šanse nego eksperti u tom polju), manjka moći izračunljivosti (ovaj projekat je radjen na procesoru koji relativno nije toliko moćan, mnogi korsite grafičku karticu za izračunavanja da bi rešili ovaj problem ali trenutno grafiška kartica nam nije bila na raspolaganju), modela(lstm model se pokazao kao najbolji za rešavanje ovog problema međjutim da bi se dobili malo bolji rezultati od 50% (nasumicno pogađanje) ovaj model se kombinuje sa fazi logikom, takodje se dodaju neke izmene same strukture lstm ćelije).

Sve ove činjenice doprinele su tome da rezultati nisu zadovoljavajući. U nakim podacima tik iznad 50% uspešnosti mreža pogađa da li će cena da raste ili opada.

Međjutim rešavanjem problema navedenih iznad, korišćenjem drugih metoda zajedno sa LSTM mrežom ,i u nekim specijalnim slučajevima stručnjaci su prijavljivali rezultate i do 55% uspešnosti.

5. Zaključak

Neuronske mreže su efikasne metode u oblasti predviđanja na tržištu hartija od vrednosti, ali je nemoguće povezati pojedinu metodologiju sa određenim problemom. Primena neuronskih mreža je najčešća u prognoziranju cena deonica., ali mnogi autori naglašavaju značaj integracije neuronskih mreža sa drugim oblicima veštačke inteligencije. Struktura neuronskih mreža u aplikacijama na tržištima hartija od vrednosti je vrlo retko optimizovana naprednim tehnikama optimizacija, a stopa tačnosti je najčešće merilo ocenjivanja.

Doprinos neuronskih mreža je u njihovoj sposobnosti da tačno predviđaju čak i u situacijama sa neizvesnim podacima, kao i mogućnost integracije sa drugim metodama.

S druge strane, ograničenja su u nedovoljnim testovima stepena oslanjanja na rezultate, nadovoljno razvijenom dizajniranju podataka i nemogućnosti određivanja optimalne toplogije (arhitekture) za određeno problemsko područje.

Mnogi autori naglašavaju potrebu za uključivanje više podataka u modele, kao npr. druge tipove imovine, više finansijskih koeficijenata, te kvalitativnih podataka. Nadalje, česte su preporuke za korišćenje različitih vremenskih razdoblja. Vrednost deonica se često predviđa na osnovu dnevnih podataka, iako neki istraživači upotrebljavaju nedeljne i

mesečne podatke. Buduća istraživanja bi trebala da se fokusiraju na ispitivanje drugih tipova mreža koje se ređe primenjuju.

Gotovo svi istraživači naglašavaju integraciju neuronskih mreža sa drugim metodama veštačke inteligencije, kao što su npr. ekspertni sistemi, kao najbolje rešenje za rešavanje nekih nedostataka.

Tehnologija nezaustavljivo napreduje i postaje pristupačna sve većem broju učesnika na tržištu. Dugoročni investitori, ulagači u deonice, fond menadžeri i svi ostali učesnici na tržištu imaju različite pragove tolerancije na rizik i ulažu u različitim vremenskim okvirima s različitim strategijama za ulaganje. Novi alati za modelovanje tržišta zasnovani na veštačkoj inteligenciji svakom investitoru pružaju priliku za povećanje dobiti i mogu postati njegova ključna prednost nad ostalim ulagačima.

Sadržaj

1. Apstrakt	1
2. Uvod.....	1
3. RNN	3
3.1 Arhitektura	4
3.2 LSTM	7
4. Primena i implementacija.	10
5. Zaključak.	10

Literatura

- [1] Primena neurosnkih mreza u finansijama, nepoznat autor
- [2] Primena veštačkih neuronskih mreža za kratkoročno predviđanje i analizu sistema daljinskog grejanja, MILOŠ B. SIMONOVIC, doktorska disertacija
- [3] <https://www.ijsr.net/archive/v6i4/ART20172755.pdf>
- [4] http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture10.pdf
- [5] <https://github.com/dzitkowskik/StockPredictionRNN/blob/master/docs/project.pdf>
- [6] <https://www.irjet.net/archives/V5/i3/IRJET-V5I3788.pdf>