

Priebežný report

Daniel Haraksim
Katedra KKUI
FEI TUKE
Košice, Slovensko

I. VSTUPNÝ DATASET

Náš vstupný dataset reprezentuje hodnoty nasnímane z jedného špekacieho pásu pri výrobe železa, kde určitý počet stĺpcov tvorí vstupné data, ktoré vedú k našej výslednej hodnote a to hodnotu rozsevu pod 5 milimetrov. Data sa snímali v 30 minútových intervaloch

Dataset sme následne očistili od nečíselných a duplikátnych hodnôt. Pri implementovaní datasetu sme si ho rozdelili na vstupné data, čo reprezentovalo 38 stĺpcov dát a výstupné data, data, ktoré budeme predikovať a to hodnotu rozsevu pod 5 milimetrov. Následne sme ich škálovali aby sme zaistili stabilitu modelu. Vstupné a výstupné dáta sme rozsekli na dáta tréningové, validačné a testovacie a to v pomere 70:15:15.

Keďže náš model je implementovaný pomocou rozhrania Tensorflow, tak sme potrebovali naše dáta pretransformovať na tenzory.

II. NEURONOVÁ SIETĚ

Ako model neuronovej siete sme sa rozhodli implementovať Gated Recurrent Unit ďalej ako GRU. Je to druh rekurentnej neuronovej siete, podobný modelu Long-Short Term Memory(LSTM). Na rozdiel od LSTM je GRU rýchlejší a vyžaduje menej dát a parametrov [1]. Skladá sa z dvoch brán, resetovacej a aktualizacej brány. Resetovacia brána určuje ako sa bude kombinovať aktuálny input v stave s historickou pamäťou. Aktualizačná brána je zodpovedná ako veľmi by mala byť pamäť udržiavaná v uzle [2]. Daný model sme si zvolil, lebo je vhodná na data v určitom časovom priebehu, čo vyhovuje našim nasnímaným dátam v časových intervaloch a slúži ako vhodná alternatíva pre porovnanie výsledkov s LSTM modelom.

Ako náš optimalizátor sme si zvolili Adam a pracovali sme s učiacou hodnotou 0.001.

Model je sekvenciálna neurónová sieť pozostávajúca z dvoch vrstiev Gated Recurrent Unit (GRU), nasledovaných viacerými plne prepojenými vrstvami. Prvá vrstva GRU má 128 jednotiek a je nastavená na návrat sekvencií, zatiaľ čo druhá vrstva GRU má 64 jednotiek a vracia iba výstup posledného časového kroku. Obe vrstvy GRU používajú aktivačnú funkciu Rectified Linear Unit (ReLU) na zavedenie nelinearity. Po vrstvách GRU nasledujú štyri plne prepojené vrstvy s 64, 32, 16 a 8 jednotkami, pričom každá používa tiež aktivačnú funkciu ReLU. Tretia a štvrtá plne prepojená vrstva majú aplikovanú regularizáciu L2 na ich maticu jadier na minimalizáciu preučenia. Po tretej hustej vrstve je pridaná vrstva Dropout s mierou 0,1

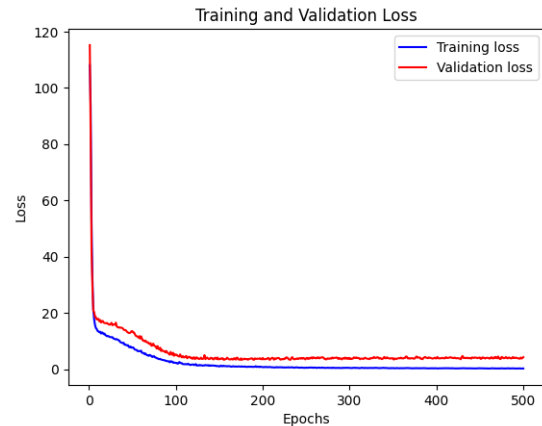


Fig. 1. Krivka strát pri tréňovaní

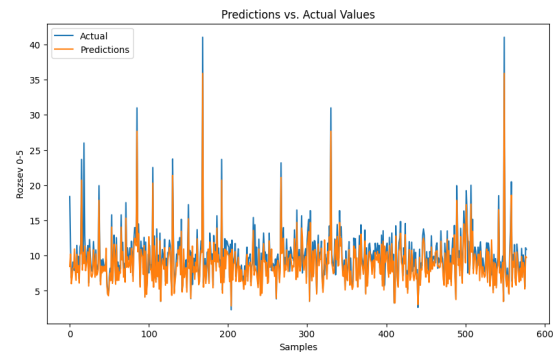


Fig. 2. Graf predikcie a reálnych hodnôt pri testovacej množine

na ďalšiu prevenciu preučenia tým, že náhodne vypúšťa 10 percent vstupných dát počas tréňovania. Posledná výstupná vrstva pozostáva z jediného hustého neurónu, ktorý predstavuje predikciu modelu pre regresné úlohy, bez špecifikovanej aktivačnej funkcie, čo naznačuje lineárnu aktiváciu (identitnú funkciu) pre predikciu spojitého výstupu. Táto architektúra je navrhnutá pre sekvencie dát s 17 príznakmi na každom časovom kroku, s cieľom predikovať jediný spojitý výstup.

III. VÝKONNOSŤ SIETE

Náš model bol tréňovaný na 500 epochoch v čase 2 minút. Na našom grafe [Fig. 1] je jasne viditeľné, že model sa našich dátach učil správne vďaka rýchlemu pádu straty a tak sa stabilizoval alebo mierne klesal.

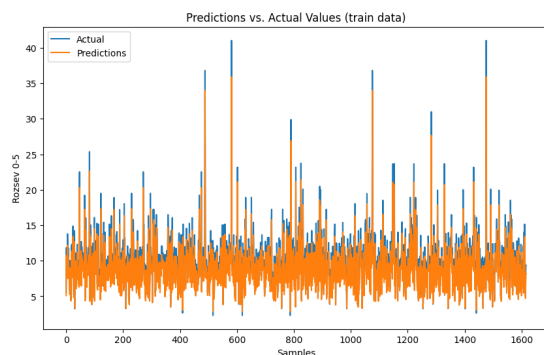


Fig. 3. Graf predikcie a reálnych hodnôt pri trénovacej množine

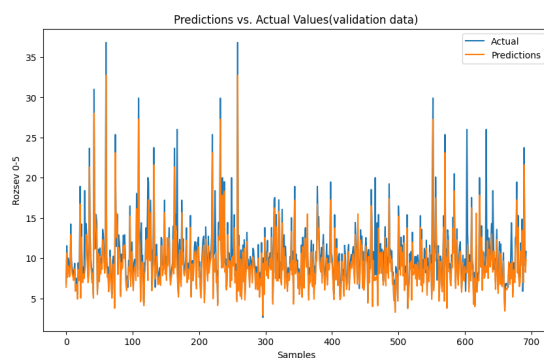


Fig. 4. Graf predikcie a reálnych hodnôt pri validačnej množine

Na grafoch porovnávania našich predikcií a reálnych hodnôt môžeme vidieť, že náš model predikoval hodnoty menšie ako reálne boli a hlavne to vidieť pri extrémnych výkyvoch. Na grafoch [Fig. 2] a [Fig. 3] je vidieť, že pri trénovaní náš model predikoval relatívne presne a tomu aj naznačovala nízka funkcia straty ale na grafe [Fig. 4] je vidieť, že funkcia straty bola vyššia a model mal väčšie nepresnosti. Našu presnosť a výsledky môžeme vidieť na našej tabuľke, kde sme využili metriky accuracy a Mean Squared Error funkciu straty.

REFERENCES

- [1] M. Pan et al., "Water Level Prediction Model Based on GRU and CNN," in IEEE Access, vol. 8, pp. 60090-60100, 2020, doi: 10.1109/ACCESS.2020.2982433. keywords: Predictive models; Rivers; Data models; Time series analysis; Adaptation models; Neural networks; Forecasting; CNN; GRU; water level prediction,
- [2] Sethia, A., Raut, P. (2019). Application of LSTM, GRU and ICA for Stock Price Prediction. In: Satapathy, S., Joshi, A. (eds) Information and Communication Technology for Intelligent Systems . Smart Innovation, Systems and Technologies, vol 107. Springer, Singapore. https://doi.org/10.1007/978-981-13-1747-7_46

	Trénovacia množina	Testovacia množina	Validačná množina
MSE-loss	0.1483	2.5585	4.4341
Accuracy	0.92131	0.82072	

TABLE I
METRIKY PRESNOSTI MODELU