

Predikcia hodnoty reozsevu aglomerátu pomocou CNN modelu

Lukáš Skopár

Technická univerzita v Košiciach
Fakulta elektrotechniky a informatiky
lukas.skopar@student.tuke.sk

Abstract—Tento článok pojednáva o využití konvolučných neurónových sietí (CNN) na predikciu reozsevu aglomerátu. Bežné prístupy k predikcii a riadeniu rozsevu čelia výzvam, ako je variabilita surovín a dynamika výrobných podmienok. CNN siete majú schopnosť efektívne analyzovať a identifikovať vzory vo vstupných dátach, čo umožňuje ich využitie na predikciu procesov ako je aglomerácia. Výskum ukazuje, že CNN modely sú vhodné pre dynamické prostredie výrobných procesov, kde sa menia podmienky a vstupné dáta. Navrhovaná je integrácia pokročilých analytických techník z oblasti strojového učenia do priemyselných prostredí s cieľom zlepšiť rozhodovanie a efektívnosť výrobných procesov. Vyhodnotenie CNN modelov v predikcii reozsevu aglomerátu prináša cenné poznatky pre optimalizáciu výrobného výkonu a kontrolu kvality.

Index Terms—Predikcia, CNN model, Umelá inteligencia, Aglomerácia

I. ÚVOD

V priemysle sa stále viac presadzujú moderné technológie, ktoré umožňujú efektívnejšie riadenie a optimalizáciu výrobných procesov. Jednou z kľúčových oblastí je predikcia hodnôt rozsevu aglomerátu, čo je kritický parameter pre kvalitu výsledného produktu. Optimalizačné ciele zahŕňajú nielen zlepšenie kvality výrobku, ale aj minimalizáciu odpadu, znížovanie nákladov a zvýšenie efektivity výrobných procesov. V kontexte predikcie hodnôt rozsevu aglomerátu je cieľom dosiahnuť presné a spoľahlivé predikcie, ktoré umožnia rýchlu reakciu na zmeny v procese a minimalizáciu nežiaducich variácií.

Tradičné metódy predikcie hodnôt reozsevu aglomerátu často spoliehajú na empirické vzťahy alebo fyzikálne modely, ktoré môžu byť obmedzené ich presnosťou a flexibilitou. V tomto kontexte môže byť použitie konvolučných neurónových sietí (CNN) revolučným prístupom. CNN modely, ktoré boli pôvodne vyvinuté pre analýzu obrazových dát, ukázali svoju silu aj v analýze časových radov a sekvencií dát. Ich schopnosť extrahovať komplexné vzory a vzťahy z dát môže poskytnúť železiarskemu priemyslu nový nástroj na presnejšiu predikciu rozsevu aglomerátu. Tieto modely sú schopné automaticky identifikovať dôležité vlastnosti a vzory vo vstupných dátach, čo umožňuje lepšie porozumenie dynamiky procesu a vytvorenie presnejších predikcií.

Hlavným zámerom článku je zhodnotiť presnosť a efektívnosť konvolučných neurónových sietí (CNN) pri

predikcii hodnôt rozsevu aglomerátu v železiarskom priemysle. Článok demonštruje, ako môže použitie týchto moderných modelov prispieť k zefektívneniu procesu výroby, optimalizácii výrobných výkonov a dosiahnutiu optimalizačných cieľov priemyslu. Vďaka ich schopnosti extrahovať zložité vzory a vzťahy z dát môže poskytnúť cenný príspevok k posunu vpred v oblasti predikcie v železiarskom priemysle.

II. DÁTA

Dáta pochádzajú z US Steel v Košiciach. Tieto dáta zahŕňujú merania zo senzorov umiestnených na špekacom páse, ktoré zaznamenávajú hodnoty rozsevu v rozsahu od 0 do 5 milimetrov. Merania boli uskutočnené v 30-minútových intervaloch počas štyroch rokov z jedného špekacieho pásu.

Hodnota rozsevu je kľúčovým parametrom, pri ktorom platí, že hodnoty menšie ako 12 milimetrov indikujú kvalitný aglomerát, zatiaľ čo hodnoty väčšie ako 12 milimetrov naznačujú nekvalitný aglomerát. Tieto dáta poskytujú dôležitý zdroj informácií pre predikciu výroby a môžu byť využité na vylepšenie procesov v železiarskom priemysle. Ich analýza a využitie v modeloch predikcie môžu prispieť k efektívnejšiemu riadeniu výrobných procesov a minimalizácii výskytu nekvalitného aglomerátu.

A. Spracovanie a analýza dát

Pri spracovaní dát sme použili dva súbory, jeden obsahujúci merania aglomerátu v intervaloch po 30 minútach a druhý s meraniami o 11:00 a 23:00 hodine. Na začiatok sme konvertovali časové údaje na dátumy a nastavili sme ich na zadaný čas, aby sme mohli ďalej pracovať s konzistentnými dátami. Následne sme prešli súbory a odstránili sme nevhodné a neplatné údaje, ktoré by mohli ovplyvniť výsledky analýzy. Týmto spôsobom sme zabezpečili, že naše dáta sú čisté a pripravené na ďalšie spracovanie a modelovanie.

Pri analýze distribúcie hodnôt rozsevu v rozsahu od 0 do 5 sme zaznamenali, že väčšina hodnôt sa koncentruje okolo hodnoty 10. Okrem toho sme zistili aj výskyt extrémnych hodnôt, ktoré sa pohybujú až k hodnote 40. Tieto extrémne hodnoty môžu naznačovať výskyt nezvyčajných udalostí alebo problémov v procese výroby, ktoré by mohli byť potenciálnymi bodmi záujmu pre ďalšiu analýzu a zlepšenie procesov. Rovnomerné rozloženie dát v pomerne stabilných intervaloch naznačuje, že merania boli uskutočnené pravidelne. Tento fakt

môže byť prospešný pri tvorbe prediktívnych modelov, pretože umožňuje modelom lepšie pochopiť rôzne aspekty výrobného procesu a predikovať hodnoty s vyššou presnosťou.

B. Náčítanie a rozdelenie vstupných dát

Po spracovaní a analýze dát boli tieto dáta normalizované metódou MinMaxScaler. Cieľom tejto normalizácie bolo zabezpečiť, že všetky hodnoty budú v rovnakom rozsahu a nebudú mať výrazný vplyv na výsledky modelovania. Dáta boli rozdelené na tréningové, validačné a testovacie množiny. Tento proces zahŕňal náhodné rozdelenie vstupných dát do tréningových a testovacích množín v určenom pomere (80% tréningových a 20% testovacích dát). Ďalej bola rozdelená tréningová množina na tréningovú a validačnú množinu, čo nám umožnilo vyhodnotiť výkon modelu počas tréningovania (70% tréningových a 30% testovacích dát). Ďalej boli preformátované tréningové, validačné a testovacie dáta na trojrozmerný tvar, ktorý je potrebný pre vstup do CNN modelu. Tento krok bol vykonaný pomocou funkcie reshape, kde bol dodržaný počet vzoriek (riadky) a počet funkcií (stĺpce) v našich dátach, a pridali sme dimenziu 1 pre počet kanálov, pretože sme pracovali s jedným kanálom vstupných dát. Týmto spôsobom sme zabezpečili, že vstupné dáta sú v správnom tvare pre použitie v CNN modeli a že každý vzor je vhodne reprezentovaný pre efektívnu analýzu a učenie sa modelu.

III. CONVOLUTIONAL NEURAL NETWORK

Konvolučné neurónové siete (CNN) sú špeciálnym typom neurónových sietí, ktoré sú široko využívané v oblasti počítačového videnia a analýzy obrazových dát. Ich sila spočíva v schopnosti efektívne extrahovať zložité vzory a vzťahy z vizuálnych dát pomocou konvolúcií a max-pooling operácií.

Využitie CNN na textových dátach je trochu odlišné od ich tradičného použitia v oblasti počítačového videnia. V prípade textových dát sa často používajú tzv. konvolučné vrstvy na detekciu vzorov vo vstupných textových sekvenciách. Tieto vrstvy môžu byť napríklad konvolučné vrstvy s oknom prechodu, ktoré sa pohybujú po vstupných dátach a aplikujú filtre na detekciu určitých vzorov.

V kontexte predikcie číselných dát, CNN môžu byť využité na predikciu hodnôt na základe vzorov a vzťahov v týchto dátach. Po prevedení dát do vhodného formátu, ako sme urobili v našom prípade, môžeme použiť CNN na naučenie sa vzorov a vzťahov v časových radoch a sekvenciách dát a následne na predikciu budúcich hodnôt. CNN môžu byť efektívne vytvorené a tréňované na týchto dátach pomocou kombinácie konvolučných vrstiev, max-pooling vrstiev a plne prepojených vrstiev.

Takéto prístupy umožňujú využitie silných stránok CNN aj v doménach mimo počítačového videnia, ako je analýza a predikcia číselných dát, ako je napríklad predikcia hodnôt rozsevu aglomerátu v železiarskom priemysle.

A. Prečo CNN

Rozhodol som sa použiť konvolučné neurónové siete (CNN) na riešenie predikcie hodnôt rozsevu aglomerátu z viacerých

dôvodov. V prvom rade, CNN sú známe svojou schopnosťou extrahovať zložité vzory a vzťahy z vizuálnych dát, ako sú obrazové dáta, a aplikovať ich na číselné dáta. S mojimi dátami potrebujem model, ktorý dokáže správne identifikovať vzory v časových radoch a predikovať hodnoty rozsevu aglomerátu.

CNN ponúkajú vysokú flexibilitu a schopnosť pracovať s rôznymi typmi dát, čo je v mojom prípade kľúčové. Ich schopnosť automaticky identifikovať dôležité vlastnosti a vzory vo vstupných dátach mi poskytuje nástroj na lepšie porozumenie dynamiky procesu v železiarskom priemysle a na vytvorenie presnejších predikcií.

Okrem toho, použitie CNN na mojich dátach je logickým krokom, pretože distribúcia hodnôt rozsevu aglomerátu vykazuje určité vzory a extrémne hodnoty, ktoré môžu byť efektívne zachytené a predikované pomocou konvolučných vrstiev.

Celkovo som si vybral CNN ako nástroj pre predikciu hodnôt rozsevu aglomerátu pre ich schopnosť efektívne pracovať s číselnými dátami a extrahovať z nich zložité vzory, čo mi umožňuje dosiahnuť presnejšie výsledky a lepšie porozumenie procesov v železiarskom priemysle.

B. Štruktúra modelu

Model začína vytvorením sekvenčného modelu, ktorý je základným kontajnerom pre rôzne typy vrstiev v Keras. Následuje pridanie vstupnej vrstvy s určeným tvarom na základe počtu funkcií (atribútov) vstupných dát.

Ďalej sú pridané konvolučné vrstvy, ktoré slúžia na detekciu vzorov vo vstupných dátach. Každá konvolučná vrstva využíva aktivačnú funkciu ReLU. Prvé dve konvolučné vrstvy majú 128 filtrov s veľkosťou jadra 5 a 3. Medzi nimi je použitá normalizačná vrstva BatchNormalization, ktorá pomáha stabilizovať tréningovanie modelu.

Po týchto konvolučných vrstvách nasleduje vrstva MaxPooling1D, ktorá znižuje dimenziu výstupu z konvolučných vrstiev na polovicu.

Následujú ďalšie štyri konvolučné vrstvy so 64 filtermi s rôznymi veľkosťami jadier od 2 po 5. medzi konvolučnými vrstvami sú použité ďalšie vrstvy MaxPooling1D, ktoré zabezpečujú zachovanie veľkosti výstupu.

Na záver nasleduje vrstva Flatten, ktorá slúži na prevedenie výstupu z konvolučných vrstiev na jednorozmerný vektor, aby mohol byť použitý ako vstup pre plne prepojené vrstvy.

Ďalej je pridaná vrstva Dropout s pravdepodobnosťou 0.1, ktorá slúži na reguláciu overfittingu modelu tým, že náhodne vypne určité neuróny počas tréningovania. Nakoniec je pridaná plne prepojená vrstva s jedným neurónom a aktiváciou ReLU, ktorá slúži na vytvorenie výstupu modelu.

IV. VYHODNOTENIE

V tejto kapitole predstavujem výsledky experimentov a ich vyhodnotenie, ktoré som získal pomocou navrhnutého modelu na predikciu veľkosti rozsevu aglomerátu. Použil som viaceré metriky na hodnotenie výkonnosti modelu.

Po vyhodnotení modelu som dosiahol priemernú presnosť (Accuracy) na úrovni 82.96%. Táto metrika ukazuje percento správne klasifikovaných príkladov zo všetkých príkladov.

Ďalšou metrikou, ktorú som použil, je Mean Absolute Error (MAE), ktorá vyjadruje priemernú hodnotu absolútnych odchýlok medzi skutočnými a predikovanými hodnotami. Táto hodnota bola zaznamenaná na úrovni 0.609.

Mean Squared Error (MSE) je ďalšia metrika, ktorú som využil na vyhodnotenie výkonu modelu. Táto metrika kvadraticizuje rozdiely medzi skutočnými a predikovanými hodnotami, čím dáva väčšiu váhu väčším odchýlkam. Model dosiahol MSE hodnotu 1.663.

R2 skóre, známe aj ako koeficient determinácie, poskytuje mieru, do akej sú vysvetlené zmeny v cieľovej premennej modelom. Hodnota R2 je v rozsahu 0 až 1, pričom hodnota blízka k 1 naznačuje, že model dobre vysvetľuje variabilitu dát. Môj model dosiahol R2 skóre na úrovni 0.887, čo poukazuje na vysokú schopnosť modelu vysvetliť variabilitu v dátach.

viesť k zlepšeniu výkonnosti a efektívnosti v železiarskom priemysle.

Ďalšie možnosti výskumu zahŕňajú skúmanie rôznych architektúr CNN, optimalizáciu parametrov modelu a rozšírenie analýzy o ďalšie faktory ovplyvňujúce veľkosť rozsevu aglomerátu.

Nadobudnuté výsledky z experimentu ukázaného v tejto práci poukazujú na priam nevyhnutné využitie neurónových sietí v priemysle pre neustály pokrok a zlepšovanie efektívnosti.

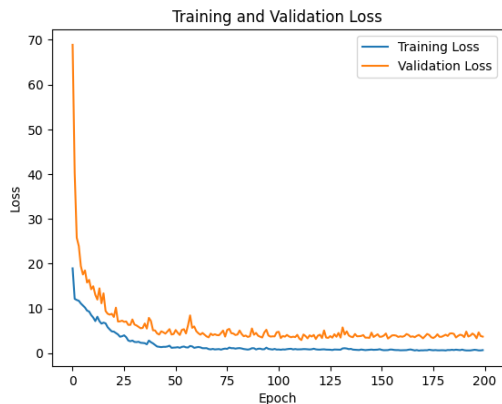


Fig. 1. Vývoj chýb pri tréovaní

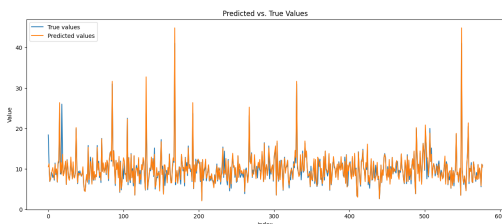


Fig. 2. Predikcia hodnôt na testovacej množine

V. ZÁVER

V tejto práci bolo preskúmané využitie konvolučných neurónových sietí (CNN) na predikciu veľkosti rozsevu aglomerátu v železiarskom priemysle. Zvolili sme CNN kvôli ich schopnosti efektívne extrahovať zložité vzory vo vstupných dátach.

Experiment preukázal účinnosť navrhnutého modelu, ktorý kombinuje viacnásobné konvolučné a max-pooling vrstvy. Tento model dosiahol vysokú presnosť pri predikcii veľkosti rozsevu aglomerátu, čo nám umožňuje lepšie porozumieť dynamike procesov v železiarskom priemysle a poskytuje nástroj na vytvorenie presnejších predikcií.

Výsledky práce naznačujú, že CNN sú vhodným nástrojom pre predikciu veľkosti rozsevu aglomerátu a ich využitie môže