

Vpeljava umetne inteligence in strojnega učenja v poslovni proces napovedovanja porabe električne energije

Vili Podgorelec¹, Sašo Karakatič¹, Grega Vrbancič¹, Špela Pečnik¹, Iztok Fister ml.¹,
Lucija Brezočnik¹, Miro Rogina², Franci Klauzner²

¹ Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko,
Maribor, Slovenija
saso.karakatic@um.si, spela.pecnik@um.si, grega.vrbancic@um.si, rok.kukovec@um.si,
vili.podogrelec@um.si

² Informatika, Informacijske storitve in inženiring, d.o.o.,
Maribor, Slovenija
miro.rogina@informatika.si, franc.klauzner@informatika.si

Sinopsis Podjetja dandanes vneto tekmujejo v zagotavljanju najboljših možnih storitev svojim strankam, pri čemer podjetja na trgu električne energije niso izjema. Glede na negotove razmere na področju zagotavljanja energetskega virov, v katerih se je znašel svet, vse večje potrebe po električni energiji in trend strme rasti cen energije je postala optimizacija poslovanja eno ključnih vprašanj celotnega sektorja. Ker umetna inteligenca (AI) in strojno učenje (ML) veljata za temeljni večnamenski tehnologiji ter za inovacijsko sredstvo z največjim potencialom preboja, si podjetja prizadevajo sprejeti te tehnologije in jih integrirati v svoje poslovne procese. Za konkretno preizkušanje možnosti uvedbe in uporabe metod AI in ML v svojem poslovnem procesu in informacijskih sistemih je podjetje Informatika informacijske storitve in inženiring d.o.o., osrednji partner v slovenskem elektroenergetskem prostoru, vzpostavilo in izvedlo pilotni projekt s Fakulteto za elektrotehniko, računalništvo in informatiko Univerze v Mariboru. Razvitih je bilo več modelov strojnega učenja za inteligentno analizo in napovedovanje porabe električne energije skupaj z nekaterimi podpornimi komponentami. Dobljeni rezultati kažejo, da lahko pravilna uporaba metod AI in ML omogoči podjetju zagotavljanje novih in naprednih storitev za svoje partnerje in različne vrste strank..

Ključne besede:

strojno učenje
umetna inteligenca
inteligentne programske rešitve
napovedovanje porabe električne energije

1 Uvod

Učinkovita obdelava poslovnih podatkov je že od nekdaj v središču elektroenergetike in z njo povezanih dejavnosti, katerim omogoča zagotavljanje varnih, zanesljivih in kakovostnih storitev ter pregledno poslovanje [1]. V času nenehne rasti obsega in podrobnosti zajetih podatkov na eni strani ter želje in potrebe po njihovem čim boljšem izkoristku na drugi smo priča izjemno hitremu razvoju informacijske tehnologije, programskih rešitev in storitev. Hkrati z rastjo količine podatkov, ki se dnevno ustvarijo, se namreč veča tudi sposobnost računalniških sistemov in pristopov za obdelavo te množice podatkov in odkrivanje novih, bolj točnih spoznanj. Pri tem so sodobne metode umetne inteligence (AI) in strojnega učenja (ML) tiste, ki snovalcem informacijskih sistemov in programskim inženirjem omogočajo razvoj novih storitev, temelječih na poglobljeni samodejni obdelavi množice zajetih podatkov. Intelligentne programske rešitve, ki temeljijo na metodah AI in algoritmih ML, lahko tako ob ustrezni implementaciji ponudijo vrsto funkcionalnosti ter nove in izboljšane poslovne rešitve, ki zagotavljajo strankam novo raven uporabniške izkušnje, ponudniku pa konkurenčno prednost.

A napredne metode ML in teorija, ki stoji za vse sposobnejšimi inteligentnimi rešitvami, so vse zahtevnejše. Medtem ko prinašajo uporabnikom vedno nove funkcionalnosti, za informatike – snovalce in razvijalce novih sistemov in rešitev – tak razvoj predstavlja resen izziv, saj jim ob svojih temeljnih in operativnih zadolžitvah ne ostane dovolj časa in priložnosti, da bi razvoju ustrezno sledili. Dejansko so izzivi še toliko večji, ker se sočasno z razvojem tehnologij pogosto spreminjajo tudi koncepti uporabe le-teh. V podjetju Informatika informacijske storitve in inženiring d.o.o. in v Laboratoriju za inteligentne sisteme na Fakulteti za elektrotehniko, računalništvo in informatiko Univerze v Mariboru smo vzpostavili pilotni projekt, da bi preizkusili možnosti za uvedbo in integracijo metod umetne inteligence in strojnega učenja v informacijski sistem in poslovne procese podjetja.

2 Napovedovanje porabe električne energije

Evropska komisija si je v Evropskem zelenem dogovoru [2] zastavila cilj, da Evropa do leta 2050 postane podnebno nevtralna. Za doseg tega cilja je potrebno resno zmanjšanje emisij toplogrednih plinov. Bistveno vlogo pri doseganju tega cilja ima raba energije, saj je npr. v letu 2016 skoraj tri četrtine svetovnih emisij (merjenih v ekvivalentih ogljikovega dioksida, CO₂-eq) povzročila poraba energije [3]. K zmanjšanju emisij poleg okoljskih vidikov hkrati vodi še ekonomska motivacija za povečanje energetske učinkovitosti [4]. Tudi zato se povečuje izraba obnovljivih virov energije – v letu 2019 se je proizvodnja električne energije iz obnovljivih virov povečala za 6% na skupno skoraj 27% delež obnovljivih virov energije v svetovni proizvodnji električne energije [5]. A za obnovljive vire energije je značilna nestanovitna proizvodnja električne energije. Ta nestanovitnost in s tem zmanjšana predvidljivost v primerjavi s konvencionalno proizvodnjo električne energije vodi do novih priložnosti za prihranke pri nabavi električne energije [6], za kar pa je potrebno čim bolj natančno predvideti obseg predvidene porabe.

Tehnično gledano zaporedje meritev porabe električne energije predstavlja časovno vrsto, ki si jo lahko predstavljamo kot dvodimenzionalni problem – na x osi je čas (kronološko zaporedje), medtem ko y os predstavlja porabo, pri čemer želimo napovedati porabo (y_i) za poljubno časovno točko v prihodnosti (x_i). Sodobne metode napovedovanja časovnih vrst temeljijo predvsem na principu zgodovinskega napovedovanja prihodnosti. Posebnost pri napovedovanju porabe električne energije je prisotnost večsmernih trendov, sezonskih in cikličnih nihanj, strukturnih prelomov, kar predstavlja določene omejitve in določa zahteve za izbiro ustreznih metod in modelov [7]. Za namen napovedovanja porabe električne energije obstaja cela vrsta metod in pristopov, od klasičnih tehnik časovnih vrst (avto regresije, modeli eksponentnega glajenja, dinamične regresije), preko tradicionalnih metod strojnega učenja pa vse do naprednih ansambelskih modelov in globokih nevronske mreže, sposobnih obravnavati nestacionarnost, heteroskedastičnost in serijsko korelacijo nestabilnih kratkoročnih podatkov [7]. Primerna vpeljava metod AI in ML je tako nujna, če želimo pri napovedovanju porabe električne energije in povezanih dejavnosti doseči zadovoljive rezultate.

3 Vpeljava metod AI in ML v poslovni proces podjetja

AI je tehnologija, ki spreminja podjetja in naše vsakdanje življenje. Je široko uporabno orodje, ki ljudem omogoča, da na novo premislijo, kako povezujemo informacije, analiziramo podatke in uporabljamo pridobljena spoznanja za boljše odločanje. AI velja za inovacijsko entiteto z največjim potencialom za preboj [8] in temeljno večnamensko tehnologijo, zlasti v povezavi z ML [9].

Z ustrezno uvedbo inteligentnih programskih rešitev in storitev lahko podjetja izboljšajo funkcionalnost izdelkov in kakovost storitev, bolje komunicirajo s strankami, racionalizirajo poslovanje ter oblikujejo predvidljive in natančne poslovne strategije. Vendar pa dandanes ni veliko podjetij, ki bi imela znanja in zmogljivosti za ustrezno obravnavo in spopadanje s celotnim procesom vpeljave AI in ML. V ta namen potrebujejo pomoč kompetentnega ponudnika znanja. Takšna pomoč običajno zajema tri glavne faze:

- faza **identifikacije** z iskanjem poslovnih primerov, v katerih je najbolj smiselno uvesti inteligentne rešitve,
- faza **načrtovanja**, v kateri se izvede raziskava o tem, kako oblikovati in razviti inteligentne metode in rešitve ter kako takšne rešitve vključiti v poslovni model podjetja, in
- faza **izvajanja**, ki vključuje pilotno izvajanje in njegovo oceno ter prenos znanja od ponudnika znanja na podjetje.

Če se podjetje odloči, da bo v celoti uporabilo preizkušene tehnologije za izbrane in morda tudi druge poslovne primere, bo moralo vzpostaviti ustrezne zmogljivosti in spretnosti. Zato je bistvenega pomena, da vse te faze spremljata svetovanje in izobraževanje.

3.1 Digitalizacija storitev z napovedovanjem porabe električne energije

Informatika je izkušeno podjetje s preko 40 let izkušenj pri razvoju IT rešitev in nujenju storitev, ki opravlja informacijske storitve po meri za vsa slovenska elektrodistribucijska podjetja. Njihove storitve vključujejo:

- sistem obračunavanja za distribucijo električne energije (za okrog 900.000 odjemalcev),
- sistem obračunavanja za izbrane dobavitelje električne energije (z več kot 140.000 strankami),
- zagotavljanje platforme za izmenjavo podatkov o trgu električne energije (enotna vstopna točka), ki jo uporabljajo vsi slovenski distributerji električne energije in vsi dobavitelji električne energije (več kot 30),
- integracijsko platformo, ki temelji na storitvenem vodilu, prek katerega je že integriranih več kot 20 različnih informacijskih sistemov, in
- aplikacije za upravljanje vseh postopkov, povezanih s potrošniki, na strani distributerja (izdajanje mnenj in smernic, soglasij za projekte, soglasij za priključitev, sporazumov o uporabi sistema itd.).

Informatika je tako podjetje, ki je že ponotranjilo procese digitalne preobrazbe, želi pa svoje poslovne procese še izboljšati z vpeljavo rešitev AI. V ta namen je z Laboratorijem za inteligentne sisteme Univerze v Mariboru, kot ponudnikom znanja z več kot 20-letnimi izkušnjami z analizo, načrtovanjem, implementacijo, integracijo in evalvacijo sistemov AI, ML in podatkovne znanosti zasnovalo in izvedlo pilotni projekt vpeljave metod AI in ML. Osnovni namen vzpostavljenega sodelovanja je bila izvedba digitalne preobrazbe dela poslovnega modela podjetja z uporabo modelov ML, ki lahko pomagajo pri pridobivanju konkurenčne prednosti pri strankah v sedanosti in prihodnosti, ter pridobivanje znanja na konkretnem projektu, da bi lahko ta znanja uporabili in vključevali v prihodnje projekte. Ker je osnovna dejavnost Informatike zagotavljanje storitev za podjetja za distribucijo električne energije, se je projekt osredotočil na analizo in napovedovanje porabe električne energije ter potencialne storitve, ki ju takšno napovedovanje omogoča. Cilj je bil tako še izgradnja pilotnega napovednega modela in operacionalizacija pilotnega modela v produkcijsko aplikacijo.

4 Vzpostavitev razvojnega okolja

Za potrebe analize podatkov, razvoj modelov strojnega učenja ter enostavnejše sodelovanje med soudeleženi smo vzpostavili razvojno okolje, temelječe na obstoječih odprtokodnih rešitvah, ki nam omogočajo prilagoditve in integracije z drugimi sistemi oz. platformami. V grobem vzpostavljeno razvojno okolje sestoji iz treh osnovnih odprtokodnih komponent – platform jupyterhub, MLflow in Apache Airflow (slika 1).



Slika 1: Konceptualni diagram zasnovane vzpostavljenega razvojnega okolja.

Vir: lasten.

Jupyterhub uporabnikom poenostavlja dostop do računskih virov brez potrebe po nameščanju razvojne programske opreme in njenim vzdrževanjem. Razvijalcem služi kot enotna vstopna točka do pred-pripravljenega razvojnega okolja z nameščeno programsko opremo in razvojnimi knjižnicami, sistemskemu administratorju pa kot orodje za nadzorovanje delovanja ter orodje za upravljanje z uporabniki in razpoložljivimi viri, ter omogoča ustrezno avtentikacijo uporabnikov kot tudi integracijo z zunanjimi avtentikacijskimi storitvami. Jupyterlab za vsakega posameznega uporabnika z uporabo t.i. spawner storitve zažene izoliran enouporabniški jupyterlab strežnik na podlagi vnaprej določenih dodeljenih sistemskih virov in vnaprej nameščene programske opreme in razvojnih knjižnic. Na tak način je omogočeno sočasno brezprekinitveno izvajanje različnih programskih kod neodvisno od ostalih uporabnikov. Zagon takšnih enouporabniških jupyterlab strežnikov je mogoč na istem gostiteljskem sistemu, na katerem je nameščen jupyterhub, ali pa na oddaljenih strežnikih. Razvojno okolje je smiselno vzpostaviti z uporabo zabojniške tehnologije Docker in orkestracijskega orodja Docker Swarm.

Odprtokodna platforma MLflow je namenjena upravljanju življenjskega cikla razvoja modelov strojnega učenja, vključno z eksperimenti in sledenjem poteku učenja, možnostjo ponovitve eksperimentov, namestitvijo napovednih modelov in njihovim centralnim registrom. Omogoča širok nabor funkcionalnosti, v našem primeru pa smo jo primarno izkoristili predvsem za sledenje napredka učenja naprednejših napovednih modelov, skupaj s sledljivostjo vrednosti različnih zagonskih hiper-parametrov. Platformo lahko uporabimo tudi kot centralni repozitorij zgrajenih napovednih modelov, katere lahko tudi namestimo v obliki spletnih storitev bodisi na lokalno ali na zunanjo infrastrukturo, kot na primer: Microsoft Azure ML, Amazon SageMaker ali Apache Spark UDF.

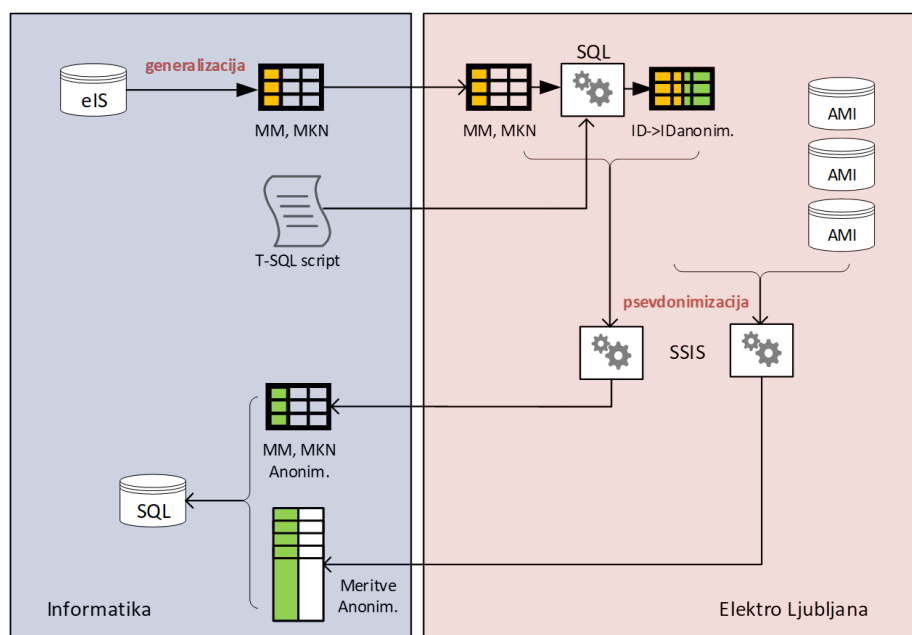
Za namen oddaljenega izvajanja dalj časa trajajočih opravil, kot je učenje kompleksnih modelov strojnega učenja, je smiselno uporabiti odprtokodno platformo Apache Airflow, ki omogoča programsko implementacijo, razporejanje in nadzorovanje delovnih tokov ter avtomatizacijo širokega nabora opravil z uporabo Python skript. Poleg skrbi za razporejanje in nadzorovanje delovnih tokov omogoča uporabnikom vpogled v izvajanje posameznih delovnih tokov z uporabo grafičnega vmesnika in integracijo z obstoječimi sistemi. Koristna funkcionalnost je še oddaljeno izvajanje jupyter zvezkov. Ker je učenje naprednejših modelov strojnega učenja časovno zahtevna naloga, lahko z uporabo Apache Airflow uporabnikom omogočimo, da želeno opravilo pošljejo v izvajanje na oddaljen sistem, sami pa nemoteno nadaljujejo z delom znotraj razvojnega okolja JupyterLab. Vpogled v napredek učenja pa je uporabnikom omogočen z uporabo prej omenjene platforme MLflow.

5 Pridobitev in pred-obdelava podatkov

Ključno vlogo pri razvoju inteligentnih rešitev z uporabo ML imajo kakovostni podatki, ki so pridobljeni na enak način in v enaki obliki, kot se pričakuje, da bodo na voljo tudi za kasnejši namen uporabe razvitih rešitev v produkcijskem okolju. V ta namen so v Informatiki poskrbeli za pridobitev podatkov o porabi električne energije na nivoju merilnih mest od Elektra Ljubljana. Definiran in vzpostavljen je bil protokol za zajem podatkov iz podatkovnih virov Elektra Ljubljane, njihovo anonimizacijo in pretvorbo v obliko, ustrezno informacijski infrastrukturi Informatike (slika 2). Na ta način je bilo pridobljeno skoraj 20 milijard meritev (19.339.546.565) za več kot pol milijona števecov.

V dejansko zajetih meritvah seveda prihaja tudi do napak, odstopanj in anomalij, v podatkovni shrambi so merilna mesta, ki več niso aktivna, na določenih merilnih mestih se lahko (tudi večkrat) zamenjajo števcji, pojavljajo se podvojene tovarniške številke (vir podatkov so tri različne baze in v njih je nekaj različnih števecov, različnih proizvajalcev in tipov, ki pa imajo enake tovarniške številke) ipd. V ta namen je bilo potrebno najprej opraviti pred-obdelavo in izločitev neustreznih podatkov. Na nivoju posameznega merilnega mesta smo izračunali različne statistične vrednosti in postavili kriterije, kateri podatki so za analizo neuporabni.

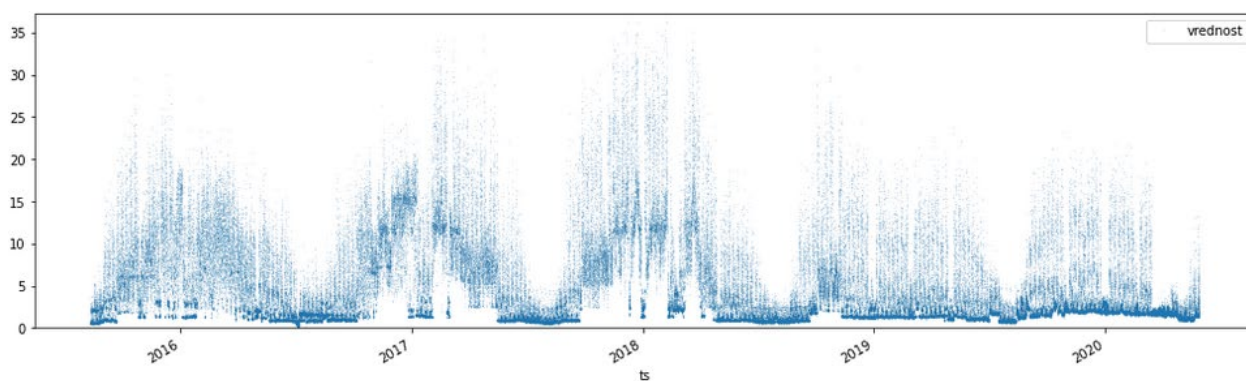
Izločili smo zapise za registre, ki predstavljajo jalovo energijo (R+, R-), tako da so ostali le podatki za prevzeto (A+, P+) in oddano (A-) delovno energijo. Izločili smo zapise za registre števecov s samimi ničelnimi vrednostmi. Za števence, ki so imeli zapisana stanja, smo vrednosti preračunali v količine energije. Izločili smo še zapise za števence s podvojenimi tovarniškimi številkami, števence s premajhnim številom zajetih meritev (npr. nedavno priključene, brez znane zgodovine porabe) ter števence, ki so imeli zapisane le vrednosti na vsako polno uro, ne pa na 15 minut.



Slika 2: Koncept pridobitve anonimiziranih podatkov o porabi.

Vir: lasten.

Nazadnje smo skušali identificirati (s primerjanjem srednje in maksimalne vrednosti ter variance) še števence, pri katerih se pojavljajo očitne napake – odbirki z nerealnimi (zelo velikimi) vrednostmi. Za namen razvoja rešitev ML je tako ostalo 12.033.896.137 meritev za 173.429 merilnih mest. Primer pred-obdelanih meritev za posamezno merilno mesto prikazuje slika 3.



Slika 3: Primer pred-obdelanih podatkov o porabi električne energije za specifično merilno mesto, ki služi kot osnova za nadaljnjo obdelavo z metodami strojnega učenja.

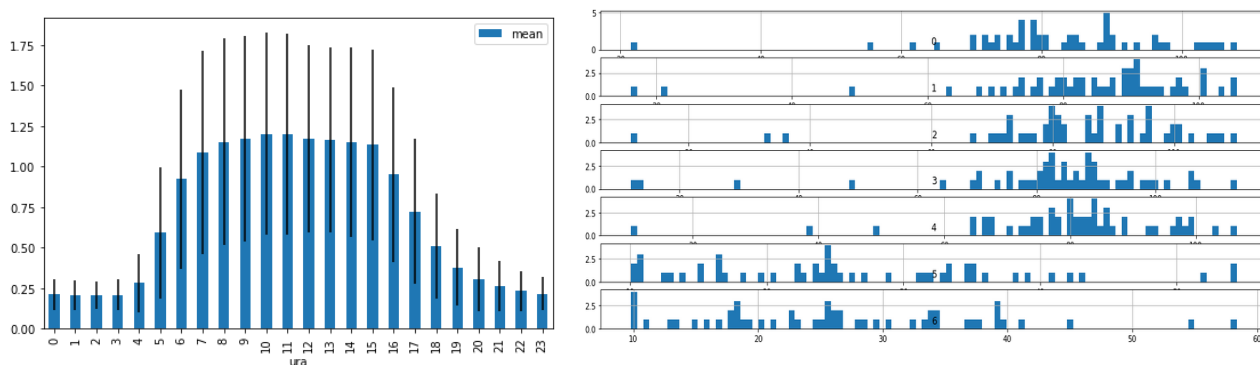
Vir: lasten.

6 Razvoj napovednih modelov s strojnimi učenjem

6.1 Inteligentna analiza podatkov – iskanje osnovnih značilnosti v meritvah porabe

Surovi podatki o porabi električne energije, zajeti v meritvah, niso neposredno uporabni za namene strojnega učenja. Zaporedje meritev porabe za posamezno merilno mesto predstavlja časovno vrsto, ki si jo lahko predstavljamo kot dvodimenzionalni problem – na x osi je čas (kronološko zaporedje), medtem ko y os predstavlja porabo. Poraba je pri tem odvisna od časa – osnovna ideja je v tem, da bi znali za poljubno časovno točko x_i (nekje v prihodnosti) izračunati (oz. napovedati) predvideno porabo y_i , ki bi se nato od dejanske porabe v tisti časovni točki kar se le da malo razlikovala. Da bi lahko to dosegli, mora obstajati povezava (korelacija) med časom in porabo, in prav to povezavo želimo poiskati. Če npr. ugotovimo, da je poraba poleti manjša kot pozimi, se bomo pri iskanju predvidene porabe za zimsko obdobje ozirali po višjih vrednostih kot pri iskanju za poletno obdobje, pri napovedovanju pa upoštevali sezono. Podobno dragocene informacije nam nudi npr. ugotovitev, da je poraba ponoči manjša kot podnevi, da je poraba višja ob delavnikih in nižja čez vikend ipd. Da bi poiskali tovrstne značilnosti, moramo čim bolj podrobno raziskati karseda velik nabor podatkov, ki so nam na voljo.

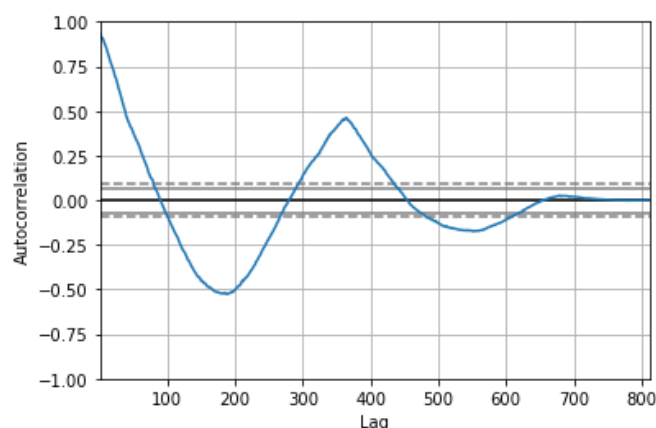
Z analizo meritev hitro ugotovimo, da je poraba precej odvisna od različnih časovnih, sezonskih (vreme), družbenih (delovni dnevi, prazniki) in ostalih zunanjih dejavnikov, ki jih bomo zato pri razvoju in gradnji napovednih modelov morali ustrezno upoštevati. Prav tako ugotovimo, da se lahko splošni vzorci porabe med različnimi merilnimi mesti zelo razlikujejo, kar pomeni, da pri napovedovanju ne bomo mogli temeljiti na enem samem, univerzalnem napovednem modelu (slika 4 prikazuje primer profila urne porabe za gospodinjstvo, ki je npr. pri poslovnem odjemalcu povsem drugačen). Po drugi strani bi razvoj in gradnja lastnih napovednih modelov za vsako merilno mesto posebej (opravka imamo z več sto tisoč porabniki, kar lahko gre v produkcijski verziji v milijone) predstavljalo daleč preveliko časovno zahtevnost problema. Sklenemo lahko, da bomo vsekakor upoštevali odkrite odvisnosti, pri čemer bomo le-te najprej izračunali za množice med seboj dovolj podobnih merilnih mest (homogene gruče odjemalcev glede na vzorce obnašanja oz. porabe) in jih šele nato prilagodili vsakemu posameznemu merilnemu mestu posebej. Na ta način dosežemo zadovoljiv in nujen kompromis med uspešnostjo napovedi ter časovno učinkovitostjo napovednega modela.



Slika 4: Primer profila urne porabe za specifično merilno mesto. Na levi sliki vidimo porabo skozi ure v dnevu, medtem ko desna slika prikazuje časovno porabo skozi različne dneve v tednu, kjer je dobro razvidna višja popoldanska poraba ob delavnikih in bolj porazdeljena urna poraba med vikendi.

Vir: lasten.

Tako na čas, potreben za učenje napovednih modelov, kot tudi na samo uspešnost napovedi ključno vpliva obseg podatkov, ki jih uporabimo za učenje modelov. Če uporabimo več podatkov (zajete meritve porabe za daljše obdobje v preteklosti), imamo na voljo več informacij za odkrivanje trendov porabe; po drugi strani pa večji obseg podatkov predstavlja večjo časovno zahtevnost za učenje napovednega modela. Zato je potrebno ugotoviti, kakšno časovno obdobje je pri razvoju napovednih modelov smiselno upoštevati. Odvisnostim, kjer nam o bodoči porabi lahko največ pove poznavanje pretekle porabe, ne glede na izhodiščni čas, rečemo avto-korelacije. Slika 5 prikazuje, kako močno je poraba v določeni časovni točki (v tem primeru gre za porabo v določenem dnevu) odvisna od porabe v prejšnjih časovnih točkah (v prejšnjih dneh). Lepo lahko vidimo, kako odvisnost s časovno oddaljenostjo pada – današnja poraba najbolj korelira z včerajšnjo, malo manj s predvčerajšnjo in tako naprej. Opazimo lahko, da avto-korelacije popolnoma opišejo sezonsko gibanje skozi letne čase, v katerih si je poraba med seboj najbolj podobna in od katerih je poraba značilno odvisna. V naslednjem letu se zadeva ponovi, a z nižjo stopnjo korelacije. V tretjem letu so zaznane stopnje korelacije že zelo nizke. Vse povedano pomeni, da je v prikazanem primeru za uspešne napovedne modele smiselno uporabiti za dve leti zgodovinskih meritev.



Slika 5: Prikaz (avto-)korelacijskih stopenj dnevne porabe električne energije, ki ponazarja, kako je vrednost porabe povezana z izmerjenimi vrednostmi v preteklosti.

Vir: lasten.

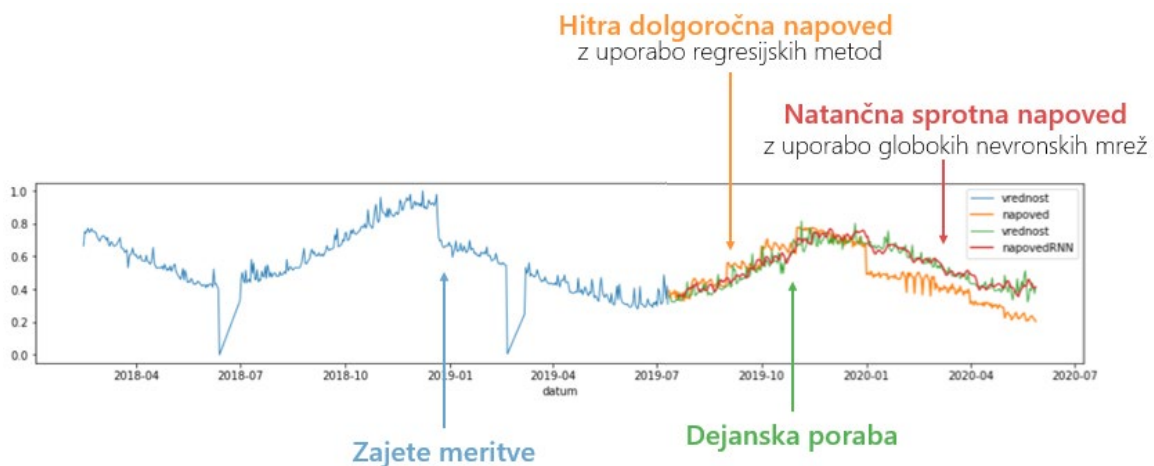
6.2 Zasnova in učenje napovednih modelov strojnega učenja

Za namen projekta smo razvili prototipno zaledno programsko rešitev za poglobljeno analizo podatkov o porabi električne energije in različne vidike napovedovanja. Osnovne razvite funkcionalnosti vključujejo:

- modele strojnega učenja za napovedovanje porabe električne energije za posameznega porabnika (dnevne, urne in 15-minutne napovedi),

- zaznavanje neobičajnih vzorcev porabe in ugotavljanje odstopanj v porabi, ter
- združevanje porabnikov električne energije na podlagi njihovih vzorcev porabe.

Temeljni del sistema je model za napovedovanje porabe električne energije. Napovedovanje porabe je bistvenega pomena za prestrukturirano okolje upravljanja energije na trgu električne energije [10]. Kljub napredku tehnologij pametnih omrežij in raziskav na področju varčevanja z energijo ostaja veliko izzivov za natančno napovedovanje proizvodnje in/ali porabe električne energije z uporabo velepodatkov ali obsežnih zbirk podatkov [11]. Za napovedovanje porabe električne energije so bili v zadnjem času uporabljeni različni pristopi, od različnih regresijskih metod [12], različnih tehnik verjetnostnega napovedovanja [13] pa vse do najnaprednejših metod ML, kot so umetne nevronske mreže [14] in zlasti metode globokega učenja [10]. Sami smo zasnovali hibridni algoritem za napovedovanje z uporabo polinomske regresije z regularizacijo za hitro posplošeno dolgoročno napovedovanje in globoko rekurentno nevronske mrežo LSTM (Long Short-Term Memory) za natančno kratkoročno napovedovanje, za katerega se je izkazalo, da zagotavlja natančne napovedi (glej primer na sliki 6). Multivariatna nevronska mreža LSTM je vsebovala niz parov slojev LSTM in Dropout s padajočo stopnjo verjetnosti osipa (dropout probability rate), končno Dense plast in optimizacijski algoritem RMSprop. Omrežje se je napajalo z dnevnimi meritvami vsaj dveh let, da bi lahko zajeli sezonske učinke. Učenje mreže je običajno potekalo, dokler ni bila dosežena stabilnost stopnje izgube (loss rate). Napovedni model porabe električne energije na nivoju posameznega odjemalca (merilnega mesta) je hkrati osnova za vse ostale razvite napovedne modele.

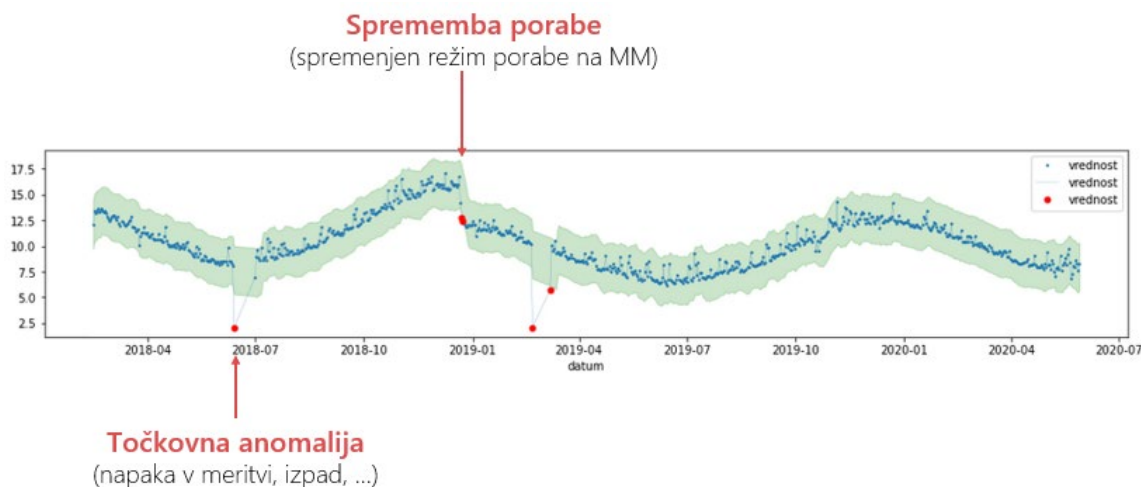


Slika 6: Primer napovedovanja porabe električne energije z razvitim pilotnim napovednim modelom.

Vir: lasten.

Drugi del sistema je komponenta za odkrivanje anomalij, ki je bila razvita za samodejno prepoznavanje nenavadne porabe električne energije in odkrivanje značilnih odstopanj od ocenjene porabe za posameznega porabnika oz. posamezno merilno mesto. V ta namen smo razvili metode in modele za zaznavanje tako točkovnih (npr. današnja poraba odstopa od vsakodnevne) kot kontekstnih nepravilnosti (npr. poraba se ni povečala/zmanjšala toliko, kot bi bilo glede na okoliščine pričakovati). Medtem ko lahko točkovne anomalije predstavljajo morebitne okvare, izpade itd., lahko kontekstne anomalije kažejo na spremenjen režim porabe električne energije določenega porabnika. Slika 7 prikazuje porabo za izbrano gospodinjstvo, kjer sta bili zaznani obe vrsti anomalij – prva ugotovljena anomalija z leve strani predstavlja točkovno anomalijo (morebitno napako pri merjenju ali izpad), medtem ko majhna skupina več točk anomalij predstavlja morebitni spremenjen režim delovanja (ker se skupna poraba od te točke zmanjša, a sledi praktično enakemu gibanju kot doslej, je lahko prišlo do prenehanja delovanja neke električne naprave ali za njeno zamenjavo z bolj varčno napravo). Odkrivanje anomalij je temeljilo na razvitim modelom napovedovanja porabe. Eksponentno zglajene napovedane vrednosti so bile uporabljene za določitev pričakovane "normalne" vrednosti, pri čemer je izračunana napaka (z upoštevanjem standardne deviacije) na učni množici predstavlja prag za "območje normalnosti". Če je dejanska meritev preseгла pričakovano območje normalnosti, je bila identificirana kot možna točkovna anomalija. Če je bilo več zaporednih meritev opredeljenih

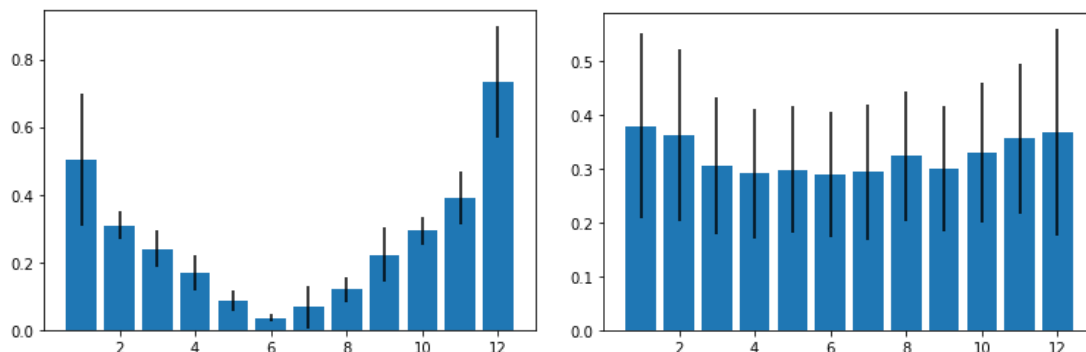
kot anomalije, je bila skupina opredeljena kot možna kontekstna anomalija, sam model odkrivanja anomalij pa se je samodejno prilagodil novemu režimu porabe.



Slika 7: Primer odkrivanja anomalij z razvitim pilotnim napovednim modelom.

Vir: lasten.

Tretji del sistema je komponenta za gručenje, namenjena analizi in segmentaciji odjemalcev v skupine podobnih porabnikov na podlagi vzorcev porabe električne energije (slika 8 prikazuje dva zelo različna profila, od katerih je vsak precej tipičen za skupino porabnikov s podobnim vedenjem). Takšno združevanje v gruče lahko omogoči npr. boljše upravljanje porabnikov in napredno statistiko porabe ali natančnejše napovedovanje, če omenimo le nekatere možnosti.

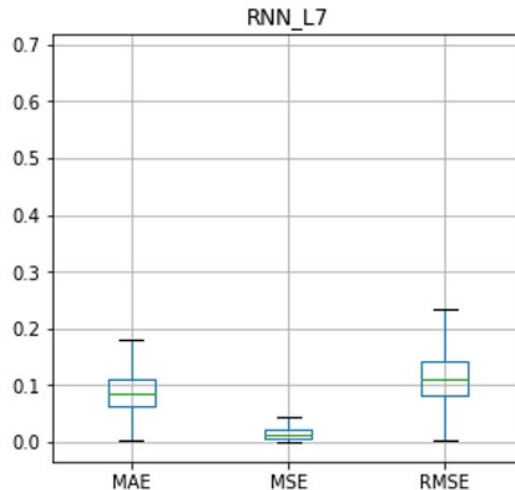


Slika 8: Primer dveh različnih tipičnih profilov porabe po mesecih, ki označujeta dve različni skupini porabnikov električne energije.

Vir: lasten.

6.3 Rezultati napovedovanja porabe električne energije

Razviti napovedni modeli so bili preizkušeni na množici meritev porabnikov električne energije. Rezultati kratkoročnega napovedovanja z uporabo rekurentnih nevronske mreže (RNN) so se izkazali za precej dobre (glej sliko 9), saj so v povprečju dosegli manj kot 10% povprečne absolutne napake (MAE) in približno 12% povprečne kvadratne napake (RMSE) – obe metriki sta izračunani na normaliziranih vrednostih meritev in tako izraženi v odstotkih. Ker je lahko poraba ob nekaterih priložnostih zelo razpršena, je standardni odklon napovedi žal precej velik (okoli $\pm 10\%$ za MAE in $\pm 12\%$ za RMSE).



Slika 9: Rezultati napovedovanja porabe električne energije z uporabo enega izmed razvitih napovednih modelov.
 Vir: lasten.

7 Namestitve in uporaba programskih storitev

Napovedne modele smo razvili s programskim jezikom Python in jih po podrobni evalvaciji izpostavili v obliki spletnih storitev kot API. Za razvoj API smo kot osnovo vzeli spletni strežnik Tornado ter nad njim implementirali spletne storitve za namen pridobivanja ter pred-obdelave podatkov, identifikacijo anomalij v podatkih, gručenje merilnih mest in napoved porabe električne energije. Za dostop do podatkov v relacijski podatkovni bazi smo se poslužili knjižnic pycopg2 ter sqlalchemy, za delo s podatki knjižnici NumPy in pandas, za potrebe implementacije metod in algoritmov identifikacije anomalij, gručenja in strojnega učenja pa smo uporabili knjižnice scikit-learn, XGBoost, scipt ter TensorFlow. Za potrebe samodejnega generiranja interaktivne dokumentacije API smo se poslužili uporabe orodja Swagger, prilagojenega za uporabo s spletnim strežnikom Tornado. Dodatno smo za pomoč pri delu z časovnimi žigi uporabili podporno knjižnico pytz. Za izgradnjo namestitvene različice API smo uporabili tehnologijo Docker, ki nam omogoča izgradnjo prenosljivih, »stateless« slik za potrebe namestitve in izvajanja API.

Po namestitvi API je njegova uporaba s programskimi klici zelo preprosta (slika 10). Preko izpostavljenih spletnih storitev smo omogočili storitve za praktično celoten proces strojnega učenja in inteligentne analize podatkov: pred-procesiranje, učenje in testiranje različnih modelov, napovedovanje porabe, odkrivanje anomalij, gručenje odjemalcev. Sistem omogoča samodejno shranjevanje in pred-pomnjenje naučenih modelov, kar izboljša časovno učinkovitost in odzivnost storitev. Poleg sklicevanja na posamezna merilna mesta, omogoča sistem tudi gradnjo in uporabo modelov za neposredno podane meritve porabe.

```
# Napovedovanje z nevronske mreže za neposredno podane meritve

interval = "day"

json_data = meritve.to_dict(orient="records") # pretvorimo DataFrame v dictionary

mm_response = requests.post(BASE_URL + f"/prediction/network/{interval}",
                             json={"ida_mm": 0, "consumption": json_data})

if mm_response.status_code == 200:
    predictions_data = mm_response.json() # json string je pretvorjen v dictionary
    train_scores = predictions_data["train_scores"]
    predictions = predictions_data["predictions"]
```

Slika 10: Primer uporabe razvitega API – klic storitve iz programskega jezika Python za napoved predvidene porabe z uporabo RNN za neposredno podane meritve.

Vir: lasten.

8 Diskusija in zaključek

V članku smo na kratko predstavili primer vpeljave metod AI in ML v poslovni proces podjetja, namenjen razvoju novih naprednih inteligentnih rešitev in storitev, temelječih na metodah, algoritmičnih in orodjih umetne inteligence. Na kratko smo pojasnili, kako lahko ponudnik znanja pomaga podjetju, ki deluje na trgu električne energije, pri sprejemanju metod ML za svoje poslovne dejavnosti in storitve. Navedli smo primer takšnega pilotnega projekta, namenjenega zagotavljanju novih naprednih storitev na podlagi učinkovitega in natančnega modela napovedovanja porabe električne energije.

Opisan pilotni projekt vpeljave metod AI in ML v poslovni proces je podjetju Informatika pomagal prepoznati, zasnovati in implementirati (trenutno v pilotnem okolju) inteligentne rešitve, ki temeljijo na umetni inteligenci in strojnem učenju ter lahko omogočijo vrsto novih in naprednih storitev njihovim poslovnim strankam in končnim uporabnikom.

Z učinkovitimi in točnimi napovednimi modeli za porabo električne energije na nivoju posameznega odjemalca oz. merilnega mesta (skupaj z drugimi komponentami) lahko Informatika različnim strankam ponudi vrsto naprednih novih storitev:

- Za distributerje in dobavitelje: optimizacija količine električne energije, optimizacija alternativnih virov, načrtovanje razvoja elektrodistribucijskega omrežja in naložb v električno omrežje, ...
- Za zunanje deležnike (npr. regulatorje): vizualizacija in analiza porabe po regijah, občinah, ..., analiza vedenjskih navad pri porabi, določitev tarifnega sistema in novih pravil za obračunavanje, ...
- Za končne uporabnike (npr. porabnike): boljše spremljanje in načrtovanje porabe, odkrivanje napak na napravah, optimalnejša poraba električne energije, ...

Pri tem je potrebno poudariti, da je bil glavni namen projekta predvsem sistematično preizkusiti, na kak način lahko podjetje načrtuje razvoj in uporabo metod strojnega učenja za zagotavljanje boljših storitev svojim strankam, ne pa razviti popolnoma optimiziran napovedni model za porabo električne energije. Čeprav razviti modeli napovedovanja kažejo obetavne rezultate, zahtevna naloga zagotavljanja tehnologije za napovedovanje porabe električne energije, ki bo primerna za produkcijsko okolje, še zdaleč ni končana. Noben posamezen model strojnega učenja ne prekaša vseh drugih modelov za vsak problem napovedovanja. Zato je izbira najboljšega algoritma odvisna od specifičnih nalog in izzivov napovedovanja. Ob tem obstaja ogromno možnosti za fino prilagajanje parametrov in nastavitev izbranih algoritmov, ki lahko bistveno vplivajo na rezultate, tako z vidika napovedne uspešnosti kot tudi glede njihove računske zahtevnosti. Poleg tega razvoj natančnih in učinkovitih modelov ML sam po sebi niti približno ni dovolj za doseganje vseh prednosti, ki jih lahko ponudi ustrezna digitalna preobrazba poslovnega procesa. Je pa vsekakor zelo dober in obetaven začetek v tej smeri, ki ga bomo zagotovo nadaljevali tudi v prihodnje.

Literatura

- [1] Zhang, Y., Huang, T., Bompard, E.F., Big data analytics in smart grids: a review, *Energy informatics*, 1(1): 1-24 (2018).
- [2] European Commission. The European Green Deal: Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions; European Commission: Brussels, Belgium, 2019.
- [3] Walther, J., Weigold, M., A systematic review on predicting and forecasting the electrical energy consumption in the manufacturing industry, *Energies*, 14(4) (2021).
- [4] Hesselbach, J.; Herrmann, C.; Detzer, R.; Martin, L.; Thiede, S.; Ludemann, B., Energy efficiency through optimised coordination of production and technical building services, In *LCE 2008: 15th CIRP International Conference on Life Cycle Engineering: Conference Proceedings*, The University of New South Wales: Sydney, Australia, p. 62 (2008).
- [5] International Energy Agency (IEA), Tracking Power. Available online: <https://www.iea.org/reports/tracking-power-2020>

- [6] Beier, J. Simulation approach towards energy flexible manufacturing systems, Sustainable Production, Life Cycle Engineering and Management; Springer International Publishing: Cham, Switzerland (2017).
- [7] Kalimoldayev, M., Drozdenko, A., Kopyk, I., Marinich, T., Abdildayeva, A., Zhukabayeva, T., Analysis of modern approaches for the prediction of electric energy consumption, Open Engineering, 10(1): 350-361 (2020).
- [8] Duan, Y., Edwards, J.S., Dwivedi, Y.K., Artificial intelligence for decision making in the era of Big Data—evolution, challenges and research agenda, International Journal of Information Management, 48: 63-71 (2019).
- [9] Lichtenthaler, U., Building blocks of successful digital transformation: Complementing technology and market issues, International Journal of Innovation and Technology Management, 17(1): 2050004 (2020).
- [10] Hafeez, G., et al.: A Novel Accurate and Fast Converging Deep Learning-Based Model for Electrical Energy Consumption Forecasting in a Smart Grid, Energies, 13(2244): 1-25 (2020). doi:10.3390/en1309224
- [11] Almalaq, A., Zhang, J.J.: Deep Learning Application: Load Forecasting in Big Data of Smart Grids, in W. Pedrycz and S.-M. Chen (eds.), Deep Learning: Algorithms and Applications, vol. 865 (New York: Springer-Verlag) 103-128. (2020)
- [12] Liu, H., Wang, Y., Wei, C., Li, J., Lin, Y., Two-Stage Short-Term Load Forecasting for Power Transformers Under Different Substation Operating Conditions, IEEE Access, 7: 161424-161436 (2019).
- [13] Nespoli, L., Medici, V., Lopaticcki, K., & Sossan, F., Hierarchical demand forecasting benchmark for the distribution grid, Electric Power Systems Research, 189: 106755 (2020).
- [14] Veeramsetty, V., Deshmukh, R., Electric power load forecasting on a 33/11 kV substation using artificial neural networks, SN Applied Sciences, 2(5): 1-10 (2020).