

UNIVERSITETI I PRISHTINËS
FAKULTETI I SHKENCAVE MATEMATIKE-NATYRORE
DEPARTAMENTI I MATEMATIKËS
PROGRAMI: MASTER NË SHKENCË KOMPJUTERIKE



TEZA MASTER

ZBATIMI I ALGORITMEVE TË TË MËSUARIT E MAKINËS PËR
PARASHIKIMIN E PRODHIMIT TË ENERGJISË NGA PANELET SOLARE

Mentori:

Prof. Ass. Dr. Artan Berisha

Kandidati:

Lyra Dobruna

Prishtinë

Nëntor, 2024

Abstrakti i zgjeruar

Të mësuarit e makinës është një ndër fushat apo disiplinat me rritjen më të shpejtë. Përdorimi i algoritmeve të të mësuarit të makinës është i gjithanshëm, duke përfshirë fusha nga më të ndryshmet si mjeksi, energjitë e rinovueshme, financa etj. Sa i përket energjive të rinovueshme, në ditët e sotme konsiderohen të jenë shumë të rëndësishme për shkak të ndikimit negative shumë të vogël në mjedis. Energjia solare është një ndër burimet më të rëndësishme për sa i përket energjive të rinovueshme, ku kapaciteti i instalimit të saj vazhdon të jetë në rritje. Për këtë arsye parashikimi i prodhimit të energjisë solare është shumë i rëndësishëm.

Në këtë punim është realizuar parashikimi i prodhimit të energjisë solare në një park solar në Kosovë bazuar në të dhënat meteorologjike dhe janë krahasuar rezultatet e fituara ndërmjet veti, si dhe me parashikimet aktuale të po këtij parku. Të dhënat rreth energjisë solare dhe parashikimeve aktuale të prodhimit të energjisë solare janë siguruar nga organet kompetente të këtij parku, kurse të dhënat meteorologjike janë siguruar nga Instituti Hidrometeorologjik i Kosovës. Të dhënat e grumbulluara janë matje të kryera në intervale të rregullta çdo orë, duke mbuluar periudhën nga data 26 janar 2022, ora 13:00, deri më 12 mars 2024, ora 08:00. Të dhënat e munguara janë trajtuar duke përdorur mesataren e variabiles përkatëse, si dhe me metodën e fshirjes në rastin e mungesës së të dhënave për një periudhë më të gjatë. Pas përpunimit të të dhënave, grupi total i tyre arrin vlerën e 17,474 rreshtave. Ndërkaq për implementimin e modeleve të parashikimit të prodhimit të energjisë solare është përdorur gjuha programuese Python, një gjuhë e nivelit të lartë, e cila ofron shumë librari me anë të të cilave mund t'i shfrytëzojmë me qindra algoritme të të mësuarit të makinës.

Për parashikimin e prodhimit të energjisë solare janë përdorur algoritme të të mësuarit të makinës, të cilat bëjnë pjesë tek mësimi i mbikëqyrur (ang. supervised learning). Algoritmet të cilat përdoren janë Përforsimi ekstrem i gradientit (ang. Extreme Gradient Boosting - XGB), Regresioni i vektorit mbështetës (ang. Support Vector Regression - SVR) dhe Pylli i rastësishëm (ang. Random Forest - RF). Këto algoritme bëjnë pjesë tek algoritmet e regresionit të bazuara në pemët e vendimmarrjes dhe regresionit.

Nga ana tjetër, hiperparametrat dhe përmirësimi i tyre luajnë një rol të rëndësishëm në përmirësimin e performancës së algoritmit. Për përmirësimin e hiperparametrave përdoren metoda të ndryshme optimizimi, ku disa prej tyre të cilat edhe do të përdoren në këtë punim janë Optimizimi Bajesian (ang. Bayesian Optimization - BO), Kërkimi i rrjetit (ang. Grid Search - GS), Kërkimi i rastësishëm (ang. Random Search - RS) dhe Optimizimi i bazuar në modele sekuenciale (ang. Sequential Model-Based Optimization - SMBO). Përdorimi i këtyre metodave të optimizimit do të na ndihmoj që të përmirësojmë performancën e algoritmeve, duke optimizuar hiperparametrat e algoritmit përkatës. Megjithëse ekziston një numër i madh i hiperparametrave, në këtë punim do të shqyrtohen disa prej tyre të cilët mund të kenë ndikim më të madh në performancën e algoritmit përkatës.

Metrikat të cilat janë shfrytëzuar në këtë punim për matjen e performancës apo vlerësimin e parashikimeve të algoritmeve janë gabimi mesatar absolut (ang. Mean Absolute Error - MAE), gabimi mesatar kuadratik (ang. Mean Square Error - MSE), rrënja e gabimit mesatar kuadratik (ang. Root Mean Square Error - RMSE) dhe koeficienti i përcaktimit (Coefficient of Determination - R^2).

Raste të ndryshme studimi janë shtjelluar në këtë punim, duke përfshirë rastin kur përdorim të gjitha fushat e grupit të të dhënave si *veçori* për parashikimin e energjisë solare, pasuar nga rastet ku bëhet optimizimi i hiperparametrave për secilin algoritëm XGB, SVR dhe RF me anë të katër metodave të optimizimit BO, GS, RS dhe SMBO. Po ashtu, janë shqyrtuar edhe disa raste duke larguar disa nga fushat e *veçorive* për parashikimin e energjisë solare.

Fillimisht këto algoritme performuan pothuajse njëllëj, ku XGB arriti saktësinë 90.07%, pasuar nga RF me 89.8% dhe SVR me 89.31%. Me inkuadrimit të metodave të optimizimit të hiperparametrave edhe performanca e algoritmeve u rrit. Algoritmi XGB performoi më së miri me metodën SMBO, duke arritur saktësinë 95.67%, metoda RF arriti rezultatin më të mirë me metodën RS, duke rezultuar me një saktësi prej 90.13%. Ndërkaq, metoda SVR arriti rezultatin më të mirë, ndonëse jo shumë më të mirë se ai aktual, me metodat GS dhe SMBO, duke rezultuar me një saktësi prej 89.41%. Ndërkaq nga ana tjetër, analizuam edhe ndikimin e rrezatimit diellor dhe presionit të ajrit tek prodhimi i energjisë solare. Së pari, fushën të cilën nuk e konsideruam në parashikim të energjisë solare ishte rrezatimi diellor, i cili njëkohësisht kishte paraqitur korrelacion të lartë me prodhimin e energjisë solare. Largimi i kësaj fushe ndikoi shumë në performancën e algoritmeve SVR dhe RF, duke e reduktuar saktësinë e tyre deri në 54%. Një ndikim në reduktimin e performancës kishte edhe tek algoritmi XGB, mirëpo në një shkallë më të vogël. Së dyti, fushën të cilën nuk e konsideruam në parashikim të prodhimit të energjisë solare ishte presioni i ajrit, i cili njëkohësisht kishte paraqitur korrelacion shumë të vogël me prodhimin e energjisë solare. Largimi i kësaj fushe pati një ndikim tejet të vogël në performancën e algoritmeve, duke e reduktuar saktësinë e tyre për një përqindje shumë të vogël. Më pas janë realizuar analiza dhe krahasime të secilit algoritëm me metodat e optimizimit të hiperparametrave, nga se mund të themi se metoda SMBO rezultoi të jetë metoda që ofroi rezultatet më të mira pothuajse në të gjitha rastet, ku në algoritmet XGB dhe SVR saktësia më e lartë u arrit me këtë metodë, kurse tek algoritmi RF rezultoi të jetë e dyta për nga saktësia me një diferencë të vogël nga pozita e parë. Po ashtu, në fund janë krahasuar rezultatet tona me parashikimet aktuale të parkut në fjalë, i cili arrin një saktësi prej 84.99%, ku krahasuar me rezultatin tonë më të mirë 95.67%, shohim se kemi arritur një saktësi rreth 10% më të madhe.

Pas zbatimit të metodave të ndryshme për parashikimin e prodhimit të energjisë nga panelet solare, mund të themi se ky punim rezulton me një performancë mjaft të mirë për parashikimin e energjisë solare bazuar në të dhënat meteorologjike. Rezultate këto që tejkaluan në saktësi edhe parashikimet aktuale nga palët e treta të realizuara nga softuer përkatës për të njejtin park solar, të cilët arritën një saktësi prej 84.99%, e cila rezulton të jetë më e dobët krahasuar me rezultatet tona të arritura në këtë punim.

Fjalët kyçe: *energjia solare, të mësuarit e makinës, regresioni, optimizimi i hiperparametrave.*