

Rezumat

Numărul și intensitatea evenimentelor meteorologice severe este în creștere, ceea ce duce la pierderi de bunuri, proprietăți și vieți umane. Prognoza meteo perfecționată, în special pentru evenimentele meteorologice severe, așteptate în mai puțin de 6 ore, cunoscută și sub numele de *nowcasting*, se dorește să ajute la atenuarea consecințelor unor astfel de evenimente. Proiectul WeaMyL își propune să îmbunătățească acuratețea prognozei pe termen scurt a vremii prin metode de învățare profundă și abordări *Big Data* capabile să gestioneze volumul mare de date meteorologice care sunt produse în mod constant. Institutele Naționale de Meteorologie sunt beneficiarii direcți ai proiectului, în timp ce publicul larg este cel mai important beneficiar indirect. Atât populația din România, cât și cea din Norvegia vor fi beneficiare, prin faptul că vor avea mai mult timp de pregătire și mai multă încredere în alertele meteorologice cu risc redus de alarmă falsă.

Echipa proiectului cuprinde cercetători de la [Universitatea Babeș-Bolyai](#) (UBB), care contribuie cu expertiza lor în învățarea automată (*machine learning* - ML). Experții meteorologi de la [Administrația Națională de Meteorologie România](#) (ANM) furnizează date și interpretări pentru România, în timp ce colegii lor de la [Institutul Meteorologic Norvegian](#) (MET) au aceleași atribuții pentru Norvegia. Echipa norvegiană cuprinde specialiști în dezvoltare software (echipa MET-IT), care sunt responsabili atât de partea de front-end a platformei WeaMyL, integrarea modulelor dintre datele meteorologice și platforma software, cât și de integrarea cu sistemele naționale de avertizare. Echipa MET include și o echipă de meteorologi operativi (echipa MET-MT) care vor testa și evalua extensiv platforma WeaMyL.

Obiectivele principale propuse pentru 2020 și 2021 au fost complet îndeplinite. Scopurile principale pentru etapa de lansare a proiectului au fost studierea abordărilor și soluțiilor existente în *nowcasting* și definirea cerințelor și arhitecturii generale pentru WeaMyL. Faza de proiect din 2021 a continuat activitatea tehnică, cu accent major pe dezvoltarea, analiza și validarea modelelor de învățare automată special adaptate.

Realizări științifice și tehnice

Prima fază a proiectului a avut loc în 2020 și a fost intitulată **Documentare, cerințe de sistem și arhitectură**. Obiectivele sale au inclus efectuarea unei analize a literaturii de cercetare existente, identificarea limitărilor din abordările și soluțiile existente în *nowcasting*, definirea cerințelor funcționale și nefuncționale pentru WeaMyL și stabilirea arhitecturii pentru platformă. Aceasta a inclus stabilirea cerințelor de utilizare și a principalelor condiții pentru buna funcționare a sistemului. A doua fază a proiectului a avut loc pe parcursul anului 2021 și a fost intitulată **Modele de învățare automată pentru nowcasting**. A doua fază a continuat activitatea desfășurată în 2020 și a adăugat noi obiective, inclusiv definirea unui model teoretic pentru *nowcasting*, dezvoltarea de modele de învățare automată scalabile special adaptate pentru *nowcasting* corect, în paralel cu validarea științifică a modelelor dezvoltate folosind analiza și interpretarea experimentală a rezultatelor. Toate activitățile planificate au fost realizate și finalizate cu succes, respectând limitele existente de timp și de buget.

În continuare, descriem activitatea desfășurată de partenerii de proiect pentru îndeplinirea obiectivelor științifice și tehnice în perioada 2020 și 2021.

Activitățile desfășurate de echipa UBB sunt rezumate în cele ce urmează. Pe parcursul anului 2020, membrii echipei au analizat metodele existente de învățare automată (ML) pentru *nowcasting*,

modele de date și arhitecturi pentru sistemele meteorologice. În timpul analizei, au fost identificate dezavantajele abordărilor existente ca punct de plecare pentru oferirea de soluții care pot depăși limitările observate. Cu această contribuție, echipa UBB a definit și structurat cerințele funcționale, ținând cont de revizuirea literaturii de specialitate împreună cu cerințele utilizatorilor finali furnizate de partenerii meteorologici. Utilizarea sistemului a fost caracterizată printr-un set de scenarii de utilizare care vor fi folosite pentru dezvoltarea sistemului. În continuare, au fost specificate modulele și interfețele principale pentru WeaMyL; această etapă a inclus definirea unui model general, teoretic, pentru *nowcasting*, bazat pe surse de date relevante, precum și pe cerințele utilizatorului final al proiectului. În ultima activitate din 2020, echipa UBB a elaborat specificații pentru datele care trebuie să fie integrate în sistem și cerințele funcționale legate de accesul la date.

Activitățile echipei UBB pentru 2021 au început cu îmbunătățirea progresivă a proiectării sistemului și a arhitecturii generale pentru WeaMyL. Aceasta a cuprins mai multe activități desfășurate în paralel. Activitățile au început prin investigarea modelelor computaționale dezvoltate dintr-o perspectivă de învățare nesupervizată, precum și din învățarea supervizată bazată pe clasificare și regresie. Apoi, echipa a investigat tehnicile de preprocesare a datelor meteorologice brute, cum ar fi pregătirea datelor pentru antrenarea metodelor, detectarea posibilelor valori aberante, precum și identificarea observațiilor meteorologice și a produselor relevante din punct de vedere computațional. În plus, aceasta a implicat investigarea tehnicilor de reducere a dimensionalității datelor, pentru a face față problemei dimensionalității. Primele metode dezvoltate au fost cele de învățare nesupervizată, care urmăreau detectarea șabloanelor în datele meteorologice și geografice preprocesate. Acestea au fost apoi aplicate pentru vizualizări și pentru a oferi unele indicații legate de predicție. În continuare, echipa UBB a investigat aplicarea metodelor de învățare supervizată, precum și dezvoltarea modelelor de învățare automată pentru *nowcasting* bazate pe arhitectura Xception. Validarea modelului Xception a fost începută în septembrie 2021. În cadrul următoarei etape a proiectului, echipa UBB are în vedere realizarea unei evaluări experimentale extinse a modelului Xception, care va include comparații și analize statistice ale rezultatelor obținute.

În paralel cu dezvoltarea metodelor ML, au fost specificate datele care trebuie integrate în sistem și cerințele funcționale legate de accesul la date. Spre sfârșitul anului 2021, echipa UBB a oferit sprijin echipei MET-IT în ceea ce privește integrarea modulului de deep learning Xception în prototipul inițial al platformei de prognoză. În plus, membrii echipei au contribuit la integrarea datelor istorice și în timp real obținute de la partenerii meteorologici în modulul de învățare profundă Xception al platformei de prognoză.

Activitățile echipei ANM în anul 2020 pot fi rezumate după cum urmează. În primul rând, echipa a revizuit tehnicile și tehnologiile consacrate și bazate pe învățare automată utilizate în *nowcasting*. Această activitate a inclus soluții actuale și emergente care sprijină meteorologii în determinarea și emiterea de avertismente meteo severe. Procesul de revizuire a fost urmat de identificarea surselor de date disponibile care au fost relevante pentru *nowcasting*. În plus, a fost furnizată o documentație structurată a tipurilor și surselor de date utilizate în prezent în prognoza meteo. Apoi, membrii echipei au furnizat cerințe ale utilizatorilor finali și cerințe nefuncționale legate de platforma WeaMyL și au definit scenarii de utilizare relevante pentru activitățile operaționale de *nowcasting*. Sprijinul a fost oferit echipelor UBB și MET-MT în timpul dezvoltării modelelor de predicție prin furnizarea de seturi de date relevante și expertiză meteorologică specifică zonei geografice.

Activitățile ANM din 2021 au început prin oferirea de feedback continuu cu privire la modelarea computațională a problemei de *nowcasting* ca sarcină de clasificare/regresie, precum și contribuind la îmbunătățirea progresivă a modelelor de calcul. Aceasta a continuat prin furnizarea de informații meteorologice cu privire la datele utilizate și rezultatele experimentelor efectuate folosind modelul Xception, atât în ceea ce privește acuratețea, cât și performanța. În plus, ANM a contribuit la

dezvoltarea Atlasului Adnotat, oferind asistență continuă în ceea ce privește integrarea datelor relevante de observare radar, satelit în prototipul inițial al băncii de date. Membrii echipei ANM au oferit feedback legat de testarea prototipului inițial Atlas.

În plus, ANM a contribuit prin oferirea accesului la baza sa de date istorice. Datele radar au fost extrase, adnotate, validate și integrate în prototipul inițial al Atlasului Adnotat. WeaMyL va fi integrat în sistemul românesc de avertizare de vreme severă, în timp ce Atlasul Adnotat va fi utilizat în procesul de previziune pentru prognozele cu mai mult de o oră înainte.

Activitățile desfășurate de echipele MET (subechipele MT și IT) este rezumată în cele ce urmează. La începutul proiectului, MET-IT a efectuat o revizuire a soluțiilor software existente pentru prognoza vremii și în special *nowcasting*. Aceasta a fost urmată de definirea surselor de date utilizate de diferiți algoritmi ML în scopuri de *nowcasting* (subechipa IT). În același timp, au fost structurate cerințele nefuncționale pentru WeaMyL și problemele de securitate privind buna funcționare a sistemului. Au fost definite măsuri cantitative și calitative pentru a valida cerințele utilizatorilor și pentru a identifica corect când acestea au fost atinse pe baza implementărilor pilot. Apoi, arhitectura generală a sistemului a fost stabilită folosind cerințele funcționale și nefuncționale identificate anterior. Membrii subechipei MT au contribuit la dezvoltarea modelelor conceptuale ale proceselor meteorologice, în special vremea severă. În acest timp, subechipa IT a implementat o componentă de achiziție de date pentru a pregăti sursele de date meteorologice ale platformei. Această componentă va furniza seturile de date necesare în continuare pentru antrenarea algoritmilor ML incluși în Platforma de Prognoză.

Activitățile din 2021 au început prin definirea conceptelor de bază și relațiilor care guvernează infrastructura platformei software de către subechipa IT. În paralel, echipa IT a început să includă în software, în mod incremental, modelele de calcul dezvoltate, în timp ce subechipa MT a procesat datele meteorologice istorice brute pentru a le aduce într-o formă adecvată pentru analize ulterioare. De asemenea, subechipa MT a asistat echipa UBB în realizarea evaluării experimentale a Xception, analiza statistică a rezultatelor obținute și compararea cu soluțiile existente. S-a oferit feedback meteorologic cu privire la performanța modelului Xception pe datele meteorologice deținute. Echipa de software de la MET a continuat cu dezvoltarea și testarea modului de integrare între WeaMyL și datele meteorologice. *Pipeline-ul* de date care va fi testat și menținut în continuare în timpul implementării componentei de achiziție de date a fost definit și configurat. În aprilie 2021, echipa IT a început dezvoltarea și testarea versiunii inițiale a Atlasului Adnotat de Observații Meteorologice, destinat în principal pentru feedback-ul timpuriu. În august 2021 a început dezvoltarea incrementală, planificată să aibă ca rezultat trei versiuni prototip – inițială, avansată și finală a platformei de prognoză. Echipa a lucrat la integrarea modului *nowcasting* prin implementarea modelului de învățare profundă Xception în Platforma de Prognoză. Componenta de achiziție de date a fost îmbunătățită urmărind feedback-ul primit de la echipele meteorologice; de asemenea, au fost integrate seturile de date meteorologice necesare pentru a construi prototipul inițial al băncii de date semantice a Atlasului Adnotat. Este planificat ca platforma WeaMyL să fie integrată în sistemul norvegian de avertizare privind vremea severă, în timp ce Atlasul Adnotat va fi utilizat în procesul de previziune pentru prognoze cu mai mult de o oră înainte.

Diseminare

În calitate de promotor al proiectului, UBB a fost responsabilă de coordonarea administrativă, științifică și tehnică a proiectului. Aceasta a implicat coordonarea comunicării și colaborării dintre parteneri, identificarea și monitorizarea riscurilor și luarea de măsuri corective atunci când este necesar. Promotorul proiectului a coordonat munca privind livrabilele și artefactele relevante pentru a se asigura că proiectul a progresat în conformitate cu constrângerile de timp și de buget pentru

îndeplinirea obiectivelor definite. Toți partenerii consorțiului au lucrat la diseminarea rezultatelor către autoritățile relevante și organizațiile științifice. În acest sens, ei au contribuit la publicarea rezultatelor activităților tehnice și științifice în reviste de înaltă calitate ce sunt public disponibile și în conferințe. Aceste eforturi s-au concretizat în publicarea a 7 lucrări științifice, dintre care 5 în reviste indexate Web of Science și 2 în conferințe Web of Science.

Una din cele mai importante activități din primul an al implementării proiectului a fost [atelierul WeADL](#), organizat în data de 28 mai 2021. Scopul său principal a fost creșterea gradului de conștientizare a comunității științifice cu privire la provocările învățării profunde, ale învățării automate sau ale cercetării științifice bazate pe inteligența artificială în domenii interdisciplinare. Accentul a fost pus pe scopurile actualului proiect -- folosirea tehnicilor de învățare profundă pentru îmbunătățirea previziunii meteorologice pe baza satelitului istoric, radar și alte produse meteorologice. Evenimentul a avut loc pe platforma Zoom, cu programul și discuțiile înregistrate și puse la dispoziție pe site-ul web. La eveniment s-au înscris peste 70 de persoane, dintre care peste 30 au participat activ.