

АНДАТПА

диссертациялық жұмысына тақырыбы:

**«МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІН ҚОЛДАНА
ОТЫРЫП, АУМАҚТЫҚ ЖОСПАРЛАУ ЖҮЙЕЛЕРІНДЕ
КЕҢІСТІКТІК ДЕРЕКТЕРДІ ТАЛДАУ ӘДІСТЕРІН ӘЗІРЛЕУ ЖӘНЕ
ЗЕРТТЕУ», 8D06102 — «Machine Learning & Data Science» білім беру
бағдарламасы бойынша философия докторы (PhD) ғылыми дәрежесін алу
үшін ұсынылған
АКЫЛБЕКОВ ОЛЖАС НАУРЫЗБАЕВИЧ**

Зерттеу жұмысының өзектілігі. Қазіргі урбанизация, халық санының өсуі және жаңа аумақтарды игеру үдерістері аумақтық жоспарлау жүйелерінің тиімділігіне жоғары талаптар қояды. Қалалар мен аймақтардың кеңістіктік дамуын басқару табиғи-климаттық, инфрақұрылымдық, әлеуметтік-экономикалық және экологиялық деректердің ірі көлемін талдауды қажет етеді. Дәстүрлі геоақпараттық әдістер көп өлшемді деректердегі жасырын заңдылықтарды ескере алмауына байланысты болжамның дәлдігін әрдайым қамтамасыз ете алмайды. Ал машиналық оқыту алгоритмдерін қолдану кеңістіктік модельдеудің, жерді қашықтықтан зондтау (ДЗЗ) деректерін интеллектуалды интерпретациялаудың және талдау үдерістерін автоматтандырудың жаңа мүмкіндіктерін ашады.

Зерттеу жұмысының мақсаты — құрылысқа жарамдылық деңгейлері бойынша аумақтарды жіктеу және сегменттеу сияқты аумақтық жоспарлау міндеттері үшін машиналық оқыту алгоритмдері негізінде кеңістіктік деректерді талдау әдістерін әзірлеу және олардың тәжірибелік тиімділігін тексеру.

Зерттеу нысаны — Алатау қаласы (Алматы облысы) бойынша цифрлық бедер үлгілері, мультиспектралды спутниктік суреттер, климаттық көрсеткіштер және көлік инфрақұрылымы жөніндегі кеңістіктік деректер.

Зерттеу пәні — әртекті рельеф пен күрделі геоақпараттық жағдайларда аумақтарды пиксельдік және аумақтық жіктеу үшін қолданылатын машиналық оқыту әдістері мен алгоритмдері.

Зерттеу міндеттері:

- Машиналық оқыту әдістерін аумақтық жоспарлауда қолданудың қолданыстағы тәсілдерін талдау;
- Көпарналы кеңістіктік деректер негізінде белгілік құрылымды әзірлеу;
- Деректерді алдын ала өңдеу: нормализация, патчтарға бөлу және аугментация жүргізу;
- CNN–MLP архитектурасына баса назар аудара отырып, машиналық оқыту модельдерін (U-Net, SegFormer, XGBoost, LightGBM) оқыту және оларды негізгі метрикалар (Accuracy, F1-score, Cohen's Kappa) бойынша салыстыру;

- Модельдердің болжамдарын эталонды схема/бас жоспармен салыстыра отырып, кросс-валидация өткізу;
- Қорытынды жарамдылық картасын құру және оның практикалық маңыздылығын бағалау.

Зерттеу әдістері. Кеңістіктік деректерді стандарттау, көпарналы белгілерді біріктіру, нейрондық және бустингтік модельдерді оқыту, кеңістіктік семантикалық сегментация, жарамдылық карталарын құру, сондай-ақ классификация сапасын бағалауға арналған статистикалық метрикалар (Accuracy, Precision, Recall, F1-score, Cohen's Kappa) қолданылды. Эксперименттер Python кітапханалары мен фреймворктары (TensorFlow/Keras, Scikit-learn, OpenCV, GDAL) және QGIS, ArcGIS ГАЖ құралдары арқылы жүзеге асырылды.

Қорғалуға ұсынылатын негізгі қағидалар:

- Кеңістіктік деректерді талдауға арналған қайталанатын есептеуіш конвейер ұсынылды: көпарналы белгілерді дайындаудан бастап жарамдылық карталарын құруға дейін;
- CNN–MLP архитектурасының тиімділігі негізделді — бұл модель шектеулі оқыту үлгілерінде жоғары дәлдік, тұрақтылық және түсініктілік көрсетті;
- Альтернативті модельдермен (U-Net, SegFormer, XGBoost, LightGBM) салыстырмалы бағалау жүргізіліп, гибридті тәсілдердің күрделі геокеңістіктік жағдайларда басым екендігі дәлелденді;
- Модель болжамдарының бас жоспардың эталондық карталарымен сәйкестігін бағалау әдістемесі әзірленді;
- Климаттық, рельефтік және инфрақұрылымдық факторларды ескерумен аумақтық жоспарлау процестерін автоматтандыру үшін машиналық оқытуды қолданудың практикалық мүмкіндігі тәжірибелік түрде расталды.

Негізгі нәтижелер. Кеңістіктік деректерді интеллектуалды талдау бойынша бағдарламалық прототип әзірленіп, жарамдылық карталарын генерациялауды толық қамтитын есептік конвейер жүзеге асырылды. Алатау ауданы деректері негізінде мультиспектралды суреттер, цифрлық бедер үлгілері және инфрақұрылымдық ақпараттар қолданылып, жүргізілген валидация нәтижелері жоғары көрсеткіштерді көрсетті:

- дәлдік (Accuracy) – 93% дейін;
- детерминация коэффициенті (R^2) – 0.92 дейін;
- F1-мөлшері – 0.91 дейін («жоғары жарамдылық» класы).

CNN–MLP архитектурасы кеңістіктік күрделілігі жоғары деректермен жұмыс істеу барысында ең жақсы және тұрақты нәтижелер көрсетті. Салыстырмалы талдау SegFormer және Vision Transformer секілді трансформерлік модельдер жақсы жалпылау қабілетіне ие екенін көрсетті, алайда CNN–MLP практикалық тұрақтылық және интерпретация жағынан басым болды, әсіресе белгіленген деректер көлемі шектеулі жағдайда.

Ғылыми жаңалығы. Кеңістіктік жарамдылықты бағалау бойынша көпкритерийлі тәсіл алғаш рет ұсынылып, кеңістіктік талдау, машиналық оқыту және ресми жоспарлау құжаттарымен салыстыру әдістері біріктірілді. Модельдер архитектурасының әсері, деректер құрылымына сезімталдығы мен тұрақтылығы жүйелі түрде зерттелді.

Практикалық маңызы. Зерттеу нәтижелері аумақтық жоспарлау жүйелерінде, жерді зонирлеу мен құрылысқа алдын ала баға беру, экологиялық бақылау және «ақылды қала» технологияларында қолдануға жарамды.

Мемлекеттік және ғылыми даму басымдықтарына сәйкестігі. Жұмыс нәтижелері Қазақстан Республикасының цифрландыру, орнықты аумақтық даму және табиғи ресурстарды тиімді пайдалану бағыттары бойынша мемлекеттік стратегияларға толық сәйкес келеді. Әдістерді ұлттық және өңірлік деңгейдегі «ақылды қала», кадастр және мониторинг жобаларында пайдалануға болады.

Автордың жеке үлесі. Докторант зерттеу мақсаттары мен міндеттерін дербес айқындап, тұжырымдамалық және техникалық шешімдерді әзірледі, бағдарламалық прототипті жүзеге асырды, модельдерді оқытып, эксперименттік валидация жүргізді және ғылыми нәтижелерді тұжырымдады. Нәтижелер негізінде ғылыми мақалалар мен тезистер дайындалды.

Апробация және жарияланымдар.

- Scopus/Web of Science базаларында индекстелетін 3 мақала (Advances in Engineering Software – Q1, IJACSA – Q3, International Journal of Artificial Intelligence – Q3);
- «ШҚТУ хабаршысы» журналында 1 мақала;
- Халықаралық конференциялар жинақтарында 1 мақала жарияланды.