

## АНДАТПА

### **6D070400 – «Есептеу технологиясы және бағдарламалық қамтамасыз ету» мамандығы бойынша «Машиналық оқыту арқылы кәсіптік бейімділікті анықтау» тақырыбындағы Таласбек Әсем Лесбекқызының кандидаттық (Ph.D.) диссертациясы**

**Зерттеудің жалпы сипаттамасы.** Бұл жұмыс кәсіби бейімділіктерді анықтау арқылы адамның жеке ерекшеліктеріне қарай болашақ мамандықты таңдауға ұсыныстар беретін қосымшаны зерттеуге және әзірлеуге арналған.

**Ғылыми жұмыстың өзектілігі.** Қазіргі таңда Қазақстандық нарықта кәсіпті анықтау жүйелері іс жүзінде жоқ. Қазіргі қоғам кәсіби шеберлікке жаңа талаптар қояды. Дегенмен, кәсіби шеберліктің жоғары деңгейі тұлғаның потенциалын толық ашуды болжайды, және оны адамның жеке ерекшеліктерін есепке алмай анықтау мүмкін емес. Ұйымдар жүргізетін көптеген сауалнамалар белгілі бір арнайы бағдарламалар бойынша тағайындау, кадрларды іріктеу үшін тұлғаның түрін жеткілікті түрде анықтап, сипаттамайды және бұл адамның белгілі бір қызметтік міндеттерді атқара алатындығы туралы сенімді нәтиже бермейді.

Мамандық бойынша кеңес беру адамдарға дана және сенімді мансаптық шешім қабылдауға көмектесу үшін арналған. Бұл шешім дұрыс өзін-өзі тануға және көптеген баламаларды мұқият қарастыруға негізделуі керек. Сонымен қатар, адамдар өз шешімдеріне қанағаттануы, таңдаған жұмысында жақсы жұмыс істеуі және болашақта мансаптық өзгерістерге немесе түзетулерге дайын болуы керек. Тұлға – балабақшада, мектепте, университетте, жанұяда, еңбек ұжымында, т.б. әртүрлі әлеуметтік жағдайларды шешу кезіндегі тұлғаның мінез-құлқы мен мінез-құлқының жиынтығы.

Адамдар немқұрайлылық пен теріс пікірге бейім, бұл олардың пайымдауларының дұрыстығына әсер етуі мүмкін. Тұлға талдауды мансап таңдау, қарым-қатынас және денсаулық туралы кеңес беру сияқты әртүрлі салалардағы бағалау ретінде қарастыруға болады. Тұлға түрі адамның оқу қабілетіне қатты әсер етеді. Мысалы, оқу үлгерімінде біз экстраверт пен интроверт адамдар арасындағы айтарлықтай айырмашылықтарды көре аламыз. Студенттердің университетті тастап кетуінің негізгі себептерінің бірі – оқу үлгерімінің төмендігі, бірақ тұлға түрі сонымен қатар интеллектуалдық қабілет, өзін-өзі бағалау, мотивация және т.б. сияқты академиялық үлгерімге әсер етеді. Зерттеулер көрсеткендей, тұлға талдауды тиімді құрал ретінде алуға болады, әсіресе университет деңгейінде академикалық көрсеткішке әсер етуі.

Twitter, Facebook, Instagram сияқты әлеуметтік желілерді қолданудың артуы автоматтандырылған түрде тұлға типін анықтау, кәсіби бейімділігін болжау – ғалымдардың өзекті мәселелерінің біріне айналдырды. Бұл зерттеу жұмыстарының негізгі теориясы: BigFive, NEO-Personality-Inventory Revised, Ten Item Personality Inventory, Myers-Briggs Type Indicator (MBTI) және т.б. Бұл саладағы жұмыстар қолданбалы бақылаудағы оқыту алгоритмдеріне негізделген.

Дегенмен, бұл мәселенің негізгі қиындығы деректер, дәлірек айтсақ, сипаттамалары болып табылады. Бұл мәселе тұлғаны болжау және жіктеу міндетін қиындатады.

**Зерттеу мақсаты.** Әртүрлі деректерді пайдаланып және машиналық оқыту әдістерін қолдана отырып, тұлғаның классификациясы негізінде адамның мамандыққа бейімділігін жоғары дәлдікпен анықтайтын қосымшаны әзірлеу.

**Зерттеу міндеттері.** Осы жұмыста алға қойылған мақсатқа сәйкес келесі міндеттер айқындалды:

- тұлғаның түрлерін және оның бейімділігін болжау және жіктеу әдістерін зерттеу және талдау;
- тұлға типтері мен кәсіптің өзара байланысын зерттеу және талдау;
- машиналық оқыту алгоритмдерін қолдану үшін әлеуметтік медиа тіркелгілерінен деректерді жинау және талдау;
- кәсіби бейімділіктерді болжау үшін эксперименттер жүргізу және модельді енгізу.

**Зерттеу объектісі.** Зерттеу жұмысында кәсіп пен тұлғаның түрін анықтаудың автоматтандырылған әдістеріне баса назар аударылады.

**Зерттеу әдістері.** Тапсырмалар теориялық және эмпирикалық зерттеулер арқылы шешілді. Зерттеуде жасанды интеллектті машиналық оқытудың классикалық теориялары мен алгоритмдерінің тұжырымдамалық ережелері, тереңдетіп оқыту үлгілері, жетекші шетелдік және отандық информатика, психология және білім беру саласындағы ғалымдардың рекомендациялық жүйелер саласындағы зерттеулері, тұлғаның жіктелуі, ықтималдықтар теориясы, математикалық статистика, сандық талдау, деректерді талдау пайдаланылды.

**Жұмыстың ғылыми жаңалығы.** Диссертациялық жұмыстың жаңалығы адамның психологиялық ерекшеліктерін ескере отырып, кәсіптік бейімділікті анықтаудың автоматтандырылған әдісін әзірлеу. Зерттеу Instagram жазбаларын және рекурентті нейрондық желі (RNN) және конволюционды нейрондық желі (CNN) үлгілерінің комбинациясын қолдану арқылы жүзеге асырылған әртүрлі эксперименттердің нәтижелерін ұсынды.

**Қорғауға жататын ғылыми талаптар:**

- Мәліметтерді жинаудың әдістері мен алгоритмдері
- Кәсіптік бейімділікті анықтау әдістері мен алгоритмдері
- Instagram желісінен жиналған түрлі деректерге негізделген тұлғаны автоматты түрде класификациялауға арналған модельдер.
- Эксперименттер, нәтижелер және талқылаулары.

**Зерттеу нәтижелерінің практикалық маңыздылығы** Зерттеу нәтижелері кәсіптік бағдар беру қызметін, оқу үлгерімін және зерттеу нәтижелерін мектеп және университет түлектеріне әртүрлі ұсынымдар жүйесінде қолдануға мүмкіндік береді, студенттердің, қызметкерлердің, қылмыскерлердің психотипі мен бейімділігін анықтауға көмектесетін жүйелерді жақсартуға көмектеседі.

**Жарияланымдар.** 6 жұмыс жарияланды, оның ішінде:

- 2 (Q2) мақала SCOPUS базасында индексиелінген International Journal of Emerging Technologies in Learning (iJET) журналында жарияланған, перцентиль-66;
- Қазақстан Республикасы Білім және ғылым министрлігінің Жоғары аттестаттау комиссиясының талаптарына сәйкес келетін 1 мақала;
- Халықаралық конференциялар материалдарында 3 баяндама.

**Диссертацияның құрылымы мен көлемі.** Диссертация көлемі 82 бет. Нормативтік сілтемелерден, анықтамалардан, қысқартулар тізімінен, кіріспеден, төрт негізгі тараудан, қорытындыдан, пайдаланылған әдебиеттер мен қосымшалардан тұрады. Диссертацияда 26 кесте, 36 сурет бар. Пайдаланылған әдебиеттер тізімі 104 атаудан тұрады.

**Бірінші тарауда** кіріспе және зерттеу жұмысының мәселелері мен мазмұны сипатталған.

**Екінші тарауда** бар жұмыстарға шолу жасалып, кәсіптік бағдар беру және жеке бейімділіктерді анықтау әдістері сипатталған.

Біріншіден, тарауда MBTI теориясы қарастырылады, ол адамды он алты тұлға түріне жіктейді. Теорияға жұппен біріктірілген сегіз көрсеткіш кіреді: экстраверсия (E) - интроверсия (I), сезім (S) - интуиция (N), ойлау (T) - сезім (F), пайымдау (J) - қабылдау (P). Әрбір тұлға түрінің функционалдық портреті ретімен сипатталады, нәтижесінде белгілі бір кәсіпке бейімділік пайда болады.

Екіншіден, тарауда тұлға типтерін болжау және жіктеу тапсырмаларында қолданылатын тәсілдер қарастырылады. Тарау әдістерді төрт санатқа бөледі: бақыланатын, бақыланбайтын, ішінара басқарылатын және терең оқыту әдістері. Әр санаттағы заманауи зерттеу жұмыстарының әдеби шолуын сипаттады. Әр санат бойынша жұмыстардың нәтижелерін салыстыру және талдау беріледі.

**Үшінші тарауда** барлық қолданбалы тәжірибелер бір-бірлеп сипатталады және олардың шектеулері талданады. Осы зерттеу жұмысындағы бірінші эксперимент k-means clustering алгоритмін пайдалана отырып, тұлғаны жіктеуге бағытталған. Ол үш бөлімнен тұрады: деректерді жинау, деректерді дайындау және гиперпараметрлерді баптау арқылы k-means кластерлеу моделін жүзеге асыру. Деректерді жинау үшін автоматтандырылған Google форма пайдаланылды. Үш кластер таңдалған кезде инерция 700 болды. Инерция қасиеті үлгілерден ең жақын кластер орталығына дейінгі қашықтықтардың квадраттарының қосындысы болып табылады. Инерция неғұрлым аз болса, жұмыс соғұрлым жақсы болады. Гиперпараметрлерді баптағаннан кейін және кластерлер санын 16-ға өзгерткеннен кейін біздің инерция параметрі 107-ге дейін төмендетілді. Жаттығу және гиперпараметрді баптау қадамдарынан кейінгі келесі қадам ретінде модель кейбір кездейсоқ кірістерді беру арқылы сынақтан өтті. Нәтижесінде ол өзі жататын кластердің индексін анықтайды.

№ 2 эксперименттің нәтижелері көрсеткендей, адамдар, әдетте, ана тілінде жүргізілсе, сауалнамаға қатысады.

№3 экспериментте Naive Bayes, XGBoost және Recurrent Neural Network модельдері зерттелді. Классикалық алгоритм мен терең оқыту алгоритмінің

сипаттамалары салыстырылды. Әрбір үлгінің сипаттамаларын сипаттау үшін дәлдік сияқты бағалау көрсеткіштері есептелді.

№4 эксперимент нәтижелері CNN кескіндерде жоғары өнімділік беретінін көрсетеді, өйткені оның «детекторлары» бар. Машинамен оқытумен байланысты барлық жұмыстар сияқты, бұл жұмыстың шектелуі деректердің болмауы болып табылады. Тағы бір айта кететін жайт, CNN нәтижелерді береді, бірақ бізге сипаттамалар мен олардың болжамға қалай әсер ететінін талдауды қамтамасыз етпейді.

№5 эксперимент Google Colab және машиналық оқыту кітапханаларын импорттау арқылы Google кластерлеріндегі Apache Spark жүйесінде орындалды. Байес үлгісінің жылдамдығы мен тиімділігін талдау үшін №3 экспериментінің бөлігі PySpark-та MapReduce парадигмасын пайдаланып қайта жазылды, мұнда белгілер кілттер және хабарламалар мәндер болып табылады.

№6 экспериментте қазақ тіліндегі мәтін пайдаланылды. Үлгі ретінде Сүлейман Демирел университетінің информатика факультетінің 450 бакалавры алынды және LSTM моделі ұсынылды.

**Төртінші тарауда** кәсіптік бейімділікті анықтайтын қосымшаның ұсынылған әдісінің бүкіл архитектурасы сипатталған. Қосымшада MBTI тұлғасының түрін және оның кәсіпке бейімділігін анықтаудың екі режимі бар. Бірінші режим мәтіндік деректерге негізделген, ал екіншісі кескін деректерін жүктеу арқылы болжауға негізделген. Бөлімде кадамдық иллюстрациямен Android қосымшасының экрандары ұсынылған.

**Бесінші тарауда** тестілеу кезеңдері, эксперимент нәтижелері және олардың өздерімен және ғылыми жұмыстың тақырыбы аясындағы қолданыстағы жұмыстармен салыстыру берілген.

**Қорытындыда** зерттеу жұмысының талдауы мен нәтижелері және оның алдағы бағыттары талқыланады. Бұл жұмыс тұлғаның қасиеттеріне негізделген кәсіптік бейімділіктерді анықтауға ықпал етеді және осы диссертацияның нәтижелерін тұлғаны тану, табиғи тілді өңдеу және жалпы білім берудегі информатика салаларындағы одан әрі зерттеу жұмыстарына қолдануға болады.