«8D06301 – Ақпараттық қауіпсіздік жүйелері» білім беру бағдарламасы бойынша философия докторы (PhD) дәрежесін алу үшін ұсынылған Сағынай Мөлдірдің «Қазақ тіліндегі деструктивті веб-контентті классификациялауға арналған тереңдетіп оқыту үлгілерін жасау» тақырыбындағы диссертациялық жұмысына

РЕЦЕНЗЕНТТІҢ ЖАЗБАША ПІКІРІ

р/н №	Өлшемшарттар	Өлшемшарттарға сәйкестігі (жауап нұсқаларының бірін сызу)	Ресми рецензенттің ұстанымына негіздеме (ескертуді курсивпен көрсету)
1.	Диссертация тақырыбының (бекіту күніне) ғылымның даму бағыттарына және/немесе мемлекеттік бағдарламаларға сәйкес болуы	1.1 Ғылымды дамытудың басым бағыттарына және/немесе мемлекеттік бағдарламаларға сәйкестігі:	Диссертация қазақ тіліндегі мәтіндерді интеллектуалды өңдеу және ақпараттық қауіпсіздік міндеттеріне арналған. Диссертация 2023-2025 жылдарға арналған ғылыми және (немесе) ғылыми-техникалық жобалар бойынша жас ғалымдарға гранттық қаржыландыруға арналған жоба аясында орындалған. Жобаның атауы мен нөмірі: AP19576868, «Жастар экстремизмін анықтау және заманауи ақпараттық кеңістікте жастардың қауіпсіздігін қамтамасыз етуге арналған модельдер мен әдістерді әзірлеу». Жобаның басым бағытының атауы: Ұлттық қауіпсіздік және қорғаныс; мамандандырылған бағыты: қолданбалы ғылыми зерттеулер; пәндік бағыты: терроризм мен экстремизмге қарсы іс-қимыл.
		1) диссертация мемлекет бюджетінен қаржыландырылатын жобаның немесе нысаналы бағдарламаның аясында орындалған (жобаның немесе бағдарламаның атауы мен нөмірін көрсету);	
		2) диссертация басқа мемлекеттік бағдарлама аясында орындалған (бағдарламаның атауын көрсету);	
		3) диссертация Қазақстан Республикасының Үкіметі жанындағы Жоғары ғылыми-техникалық комиссия бекіткен ғылым дамуының басым бағытына сәйкес (бағытын көрсету) келеді.	
2.	Ғылым үшін маңыздылығы	Жұмыс ғылымға елеулі үлесін қосады/қоспайды, ал оның маңыздылығы ашылған/ашылмаған.	Бұл диссертациялық жұмыс ғылымға елеулі үлесін қосады және оның маңыздылығы ашылған. Жұмыстың ғылыми үлесі қазақ тіліндегі деструктивті мәтіндерді контекстік, морфологиялық және пәндік семантикалық ақпаратты бірлестіріп тану тәсілін әзірлеуден көрінеді. Автор төрт класты корпус негізінде MFA модулі мен онтологиялық-графтық белгілердің үлесін зерттеген. Нәтижелер аз ресурсты

			агглютинативті тілдердегі көпкласты мәтіндік классификациялау әдістерін дамыту үшін маңызды.
3.	Өзі жазу принципі	Өзі жазу деңгейі: 1) жоғары; 2) орташа; 3) төмен; 4) өзі жазбаған.	Докторанттың дербес ғылыми үлесі зерттеу есебін қоюдан, корпус пен аннотациялау рәсімін ұйымдастырудан, морфологиялық белгілер құрамын таңдаудан, онтологиялық белгілерді модельге енгізуден және эксперимент нәтижелерін түсіндіруден байқалады. Диссертацияда автор ұсынған шешімдер пайдаланылған базалық модельдерден ажыратылып сипатталған. Әдістерді салыстыру мен қорытындыларды тұжырымдау жұмыстың дербес зерттеу сипатын көрсетеді. Диссертацияның негізгі нәтижелері халықаралық конференцияларда және жоғары импакт-факторлы журналдарда жарияланып, ғылыми қауымдастық тарапынан мойындалған.
4.	Ішкі бірлік принципі	4.1 Диссертация өзектілігінің негіздемесі: 1) негізделген; 2) ішінара негізделген; 3) негізделмеген.	Диссертацияның өзектілігі толық негізделген. Өзектілік қазақ тілінің агглютинативті құрылымы, әлеуметтік желілердегі бейресми тіл және деструктивті мағынаның контекске тәуелділігі арқылы көрініс табады. Арнайы аннотацияланған ресурстардың шектеулілігі мен жасырын агрессияны танудың күрделілігі зерттеу міндетін айқындайды. Кіріспе және бірінші тарау осы мәселелерді ұсынылған модельдік шешіммен байланыстырады.
		4.2 Диссертация мазмұны диссертация тақырыбын айқындайды: 1) айқындайды; 2) ішінара айқындайды; 3) айқындамайды.	Диссертация мазмұны диссертация тақырыбын толық айқындайды. Диссертация мазмұны атауында көрсетілген қазақ тіліндегі деструктивті веб-контентті классификациялау міндетін ашады. Екінші тарауда деректер мен белгілер, үшінші тарауда терең оқыту архитектуралары, төртінші тарауда олардың эксперименттік бағалануы берілген.
		4.3. Мақсаты мен міндеттері диссертация тақырыбына сәйкес келеді: 1) сәйкес келеді;	Диссертацияның мақсаты мен міндеттері диссертация тақырыбына толық сәйкес келеді. Қазақ тіліндегі деструктивті веб-контентті классификациялайтын терең оқыту модельдерін зерттеу және құру мақсаты міндеттермен

		2) ішінара сәйкес келеді; 3) сәйкес келмейді.	нақтыланған. Корпус дайындау, модельдерді зерттеу, морфологиялық және семантикалық белгілерді әзірлеу, гибридті архитектура құру және оны сынау бір ғылыми міндеттің өзара байланысты кезеңдерін құрайды. Барлық қойылған міндеттер дәйекті түрде орындалған.
		4.4. Диссертацияның барлық бөлімдері мен ережелері логикалық байланысқан: 1) толық байланысқан; 2) ішінара байланысқан; 3) байланыс жоқ.	Диссертацияның барлық бөлімдері мен құрылымы логикалық тұрғыдан толық байланысқан. Жұмыс теориялық талдаудан деректерді дайындауға, одан белгілер мен архитектураны әзірлеуге және эксперименттік тексеруге дәйекті түрде өтеді. Қорғалатын ережелер екінші және үшінші тараулардағы әдістемелік шешімдерге сүйеніп, төртінші тарауда сандық нәтижелермен бағаланады. Жалпы қорытынды қойылған міндеттерге сәйкес құрылған. Бұл диссертациялық жұмыстың ғылыми жүйелілігін және логикалық дәйектілігін айқын көрсетеді.
		4.5 Автор ұсынған жаңа шешімдер (қағидаттар, әдістер) дәлелденіп, бұрыннан белгілі шешімдермен салыстырылып бағаланған: 1) сыни талдау бар; 2) талдау ішінара жүргізілген; 3) талдау өз пікіріне емес, басқа авторлардың сілтемелеріне негізделген; 4) талдау жоқ.	Диссертацияда автор ұсынған жаңа шешімдер (қағидаттар, әдістер) толық дәлелденіп, бұрыннан белгілі тәсілдермен салыстырылып бағаланған және олардың сыни талдауы берілген. Автор базалық KazBERT моделін MFA, онтологиялық-графтық белгілер және баламалы тізбектік модульдер қосылған нұсқалармен салыстырылып, сандық көрсеткіштер арқылы дәлелденген. Қазақ тіліндегі деструктивті мәтіндер корпусын құру тәжірибесі аталып, тілдік ресурстар тапшылығы жағдайында бұл шешімнің ғылыми және практикалық маңызы негізделген.
5.	Ғылыми жаңашылдық принципі	5.1 Ғылыми нәтижелер мен ережелер жаңа ма? 1) толығымен жаңа; 2) ішінара жаңа (25-75% жаңа); 3) жаңа емес (жаңасы 25%-дан кем).	Диссертациялық жұмыстың барысында алынған ғылыми нәтижелер мен қағидаттар толығымен жаңа болып табылады. Жаңашылдық автор ұсынған нақты нәтижелерге қатысты бағаланады: қазақ тіліндегі төрт класты корпус, контенттің түрі, ашықтығы және нысанасын ескеретін семантикалық белгілер, оларды морфологиялық ақпаратпен және KazBERT көріністерімен біріктіру әдісі. Белгілі нейрондық компоненттер негізінде осы зерттеу міндетіне

			бейімделген, экспериментпен тексерілген авторлық архитектура қалыптастырылған.
		5.2 Диссертацияның қорытындылары жаңа ма?	Диссертацияның қорытындылары толығымен жаңа болып табылады. Диссертация жаңа эмпирикалық қорытындылар ұсынады. Морфологиялық ақпаратты, онтологиялық белгілерді енгізу бағалау параметрлерін арттырған. Онтологиялық белгілері бар нұсқаға BiLSTM қосылғанда қосымша өсім алған. Бұл тұжырымдар зерттелген корпус пен эксперименттік хаттама негізінде алынған. Қорытындылай келе, диссертацияда ұсынылған тәсілдер ғылыми тұрғыда негізделіп қана қоймай, ақпараттық қауіпсіздік пен мәтінді өңдеу күрес саласына практикалық үлес косты.
		1) толығымен жаңа;	
		2) ішінара жаңа (25-75% жаңа);	
		3) жаңа емес (жаңасы 25%-дан кем).	
		5.3 Техникалық, технологиялық, экономикалық немесе басқару шешімдері жаңа және негізделген бе?	Диссертациялық жұмыста ұсынылған техникалық және технологиялық шешімдер толығымен жаңа. Шешімнің жаңалығы белгілі KazBERT, графтық нейрондық желі және BiLSTM компоненттерін қазақ тілінің морфологиялық белгілерімен және пәндік онтологиямен біріктіруінде. Оның негізділігі модель нұсқаларын салыстыру арқылы көрсетілген.
		1) толығымен жаңа;	
		2) ішінара жаңа (25-75% жаңа);	
		3) жаңа емес (жаңасы 25%-дан кем).	
6.	Негізгі қорытындылардың негізділігі	Барлық негізгі қорытындылар ғылыми тұрғыдан қарағанда ауқымды дәлелдемелерде негізделген/негізделмеген (qualitative research (куолитатив ресеч) және өнер және гуманитарлық ғылымдар бойынша даярлық бағыттары үшін).	Диссертациялық жұмыстың негізгі қорытындылары ғылыми тұрғыдан қарағанда толық негізделген. Негізгі қорытындылар тәжірибелерге, модельдік конфигурацияны салыстыруға, бес іске қосу бойынша орташа мәндер мен стандарттық ауытқуларға сүйенеді. Accuracy, Macro-F1 және MCC нәтижелері бір бағыттағы жақсаруды көрсетеді. Қателесу матрицалары және McNemar тексеруі сандық бағалауды толықтырады.
7.	Қорғауға шығарылған негізгі ережелер	Әрбір ереже бойынша келесі сұрақтарға жеке жауап беру қажет:	Қорғауға шығарылған негізгі ережелер дәлелденді: 1-ереже. Корпус құру және контекстік, морфологиялық, онтологиялық белгілерді біріктіру 2-3-тарауларда сипатталған. 4.3-кестеде Accuracy 88,015%-дан 90,492%-ға, Macro-F1 88,086%-дан 90,567%-ға өскен.
		7.1 Ереже дәлелденді ме?	
		1) дәлелденді;	
		2) шамамен дәлелденді;	
		3) шамамен дәлелденбеді;	
		4) дәлелденбеді;	

		5) бұл тұжырымда ереженің дәлелденгенін тексеру мүмкін емес.	2-ереже. BiLSTM енгізілген толық модель Accuracy 91,045% және Macro-F1 91,103% нәтижесіне жеткен. Қате үлесінің салыстырмалы төмендеуі шамамен 25,28%. Бұл ережелер (7.1) дәлелденді, (7.2) тривиалды емес, (7.3) жаңа және (7.4) қолдану деңгейі кең, (7.5) мақалаларда дәлелденген. Диссертациялық жұмыстың нәтижелері 10 ғылыми басылымда жарияланды. 1 мақала Scopus базасында индекстелетін басылымда, 3 мақала Scopus базасында индекстелетін конференцияларда, 2 мақала ҚР БҒМ Білім және ғылым саласындағы бақылау комитеті ұсынған журналдарында жарық көрді. 4 мақала халықаралық конференциялар материалдарында жарияланып, ҚР авторлық құқықпен қорғалатын объектілерге құқықтарды мемлекеттік тізілімге мәліметтерді енгізу туралы 1 куәлік алынды.
		7.2 Тривиалды ма?	
		1) ия;	
		2) жоқ;	
		3) бұл тұжырымда ереженің тривиалды екенін тексеру мүмкін емес.	
		7.3 Жаңа ма?	
		1) ия;	
		2) жоқ;	
		3) бұл тұжырымда ереженің жаңашылдығын тексеру мүмкін емес.	
		7.4 Қолдану деңгейі:	
		1) тар;	
		2) орташа;	
		3) кең	
		4) бұл тұжырымда ереженің қолдану деңгейін тексеру мүмкін емес.	
		7.5 Мақалада дәлелденген бе?	
		1) ия;	
		2) жоқ	
		3) бұл тұжырымда мақаладағы ереженің дәлелденгенін тексеру мүмкін емес.	
8.	Дәйектілік қағидаты.	8.1 Әдіснаманы таңдау – негізделген немесе әдіснама нақты жазылған:	Ия, диссертациялық жұмыста қолданылған әдістеменің таңдауы толық негізделген және әдіснама нақты жазылған. Зерттеу міндеттерін шешу үшін машиналық және терең оқыту әдістері, мәтінді классификациялау әдістері, табиғи тілді өңдеу әдістері, әлеуметтік желілерді талдау әдістері, онтологиялық классификациялау әдістері қолданылады. Әдіснама зерттеу нысанына сәйкес: мәтіннің контексті
	Дереккөздер мен ұсынылған ақпараттың дәйектілігі	1) ия;	
		2) жоқ.	

			KazBERT арқылы, грамматикалық сипаттары MFA векторы арқылы, пәндік байланыстары онтологиялық-графтық белгілер арқылы беріледі. Деректерді бөлу, модельдерді ортақ шарттарда салыстыру және бірнеше метриканы қолдану негізделген.
		8.2 Диссертация жұмысының нәтижелері компьютерлік технологияларды қолдану арқылы ғылыми зерттеулердің қазіргі заманғы әдістері мен деректерді өңдеу және интерпретациялау әдістемелерін пайдалана отырып алынған:	Иә, диссертациялық жұмыстың нәтижелері қазіргі заманғы компьютерлік технологияларды қолдану арқылы алынған. Автор ғылыми зерттеулерде заманауи әдістерді тиімді пайдаланып, алынған деректерді өңдеу және интерпретациялау тәсілдерін ұтымды қолданған. Модельдер Python, PyTorch және Transformers негізінде іске асырылған. Контекстік тілдік модель, графтық өңдеу, рекурренттік қабаттар және көпкласты бағалау метрикалары қолданылған. Зерттеуде классификация сапасымен қатар іске қосудың кездейсоқ факторларына сезімталдық қарастырылып, эксперимент нәтижелері кестелер мен қателесу матрицалары арқылы түсіндірілген.
		1) ия;	
		2) жоқ.	
		8.3 Теориялық қорытындылар, модельдер, анықталған өзара байланыстар және заңдылықтар эксперименттік зерттеулермен дәлелденген және расталған (педагогикалық ғылымдар бойынша даярлау бағыттары үшін нәтижелер педагогикалық эксперимент негізінде дәлелденеді):	Иә, диссертациялық жұмыста ұсынылған теориялық қорытындылар, модельдер мен анықталған өзара байланыстар және заңдылықтар жүргізілген эксперименттік зерттеулер арқылы толық дәлелденген және расталған. Ұсынылған белгілерді біріктірудің тиімділігі компоненттік салыстырумен расталған. Толық модельдің бес іске қосудағы Accuracy көрсеткіші 91,045 ± 0,656%, Macro-F1 көрсеткіші 91,103 ± 0,654% болған. Бұл дәлелдер қарастырылған эксперименттік жағдайда оң қорытынды жасауға негіз береді.
		1) ия;	
		2) жоқ.	
		8.4 Маңызды мәлімдемелер нақты және сенімді ғылыми әдебиеттерге сілтемелермен расталған / ішінара расталған / расталмаған.	Диссертациялық жұмыстағы маңызды мәлімдемелер нақты және сенімді ғылыми әдебиеттерге, сонымен қатар заманауи әдістемелік және тәжірибелік еңбектерге жасалған сілтемелер арқылы расталған. Тілдік модельдер, морфологиялық ақпарат, онтологиялар, графтық нейрондық желілер және классификация нәтижелерін бағалау жөніндегі негізгі

			әдістемелік тұжырымдар ғылыми еңбектерге сүйенеді. Автордың эксперименттік мәлімдемелері өз нәтижелерімен байланыстырылған. Әдебиетке сілтемелер мен зерттеудің сандық дәлелдері бірін-бірі толықтырады.
		8.5 Пайдаланылған әдебиеттер тізімі әдеби шолуға жеткілікті/жеткіліксіз.	Пайдаланылған әдебиеттер тізімі табиғи тілді өңдеу, деструктивті контентті анықтау, қазақ тілінің морфологиясы, біліммен толықтырылған модельдер және эксперименттердің тұрақтылығы мәселелерін қамтиды. Іргелі еңбектер мен кейінгі зерттеулердің қатар берілуі зерттеу есебін қоюға және әдіснамалық таңдау жасауға жеткілікті негіз қалыптастырады.
9	Практикалық құндылық қағидаты	9.1 Диссертацияның теориялық маңызы:	Иә, диссертацияның теориялық маңызы бар. Теориялық маңызы аз ресурсты агглютинативті тілдегі деструктивті мәтінді контекстік, грамматикалық және пәндік семантикалық ақпарат арқылы сипаттау тәсілін негіздеуде. Компоненттік салыстыру белгілердің қосымша үлесін бағалауға мүмкіндік береді және кейінгі модельдер үшін тексерілетін ғылыми болжамдар ұсынады.
		1) бар;	
		2) жоқ.	
		9.2 Диссертацияның практикалық маңызы бар және алынған нәтижелерді практикада қолдану мүмкіндігі жоғары:	Иә, диссертацияның практикалық маңызы бар және алынған нәтижелерді қолдану мүмкіндігі жоғары. Қалыптастырылған корпус пен модельді қазақ тіліндегі пікірлерді алдын ала талдауда, күмәнді хабарламаларды сарапшыға іріктеп ұсынуда және контент модерациясын қолдауда пайдалануға болады. 4.6-бөлімдегі веб-прототип базалық және толық модельдің бір мәтінге берген шешімдерін салыстыруға мүмкіндік береді. Бұл нәтиже қолданбалы іске асыру мүмкіндігін көрсетеді.
		1) ия;	
		2) жоқ.	
		9.3 Практикалық ұсыныстар жаңа ма?	Практикалық ұсыныстар толығымен жаңа болып табылады. Ұсыныстардың жаңалығы веб-ресурстардағы деструктивті веб-контентті классификациялауда қазақ тілінің морфологиялық ерекшеліктерін және деструктивті ұғымдар арасындағы байланыстарды ескеретін модель арқылы іске асыруында. Сондықтан практикалық ұсыныстар жаңа және зерттеу нәтижелерімен негізделген деп бағаланады.
		1) толығымен жаңа;	
		2) ішінара жаңа (25-75% жаңа);	
		3) жаңа емес (жаңасы 25%-дан кем).	
10.	Жазу және ресімдеу сапасы	Академиялық жазу сапасы:	Академиялық жазу сапасы жоғары. Диссертация академиялық талаптарға сай жазылған. Тақырыптың
		1) жоғары;	

		2) орташа; 3) орташадан төмен; 4) төмен.	мазмұны мен құрылымы жүйелі түрде баяндалып, мәтіннің ғылыми стилі сақталған. Кестелер мен суреттер түсінікті рәсімделген. Әдебиеттер тізімі халықаралық стандарттарға сәйкес рәсімделген. Құжаттың жалпы тілі түсінікті әрі ресми-ғылыми сипатта рәсімделген.
11.	Диссертацияға ескертулер	1. Корпусты дайындау кезеңдерін қысқаша алгоритм түрінде жинақтап көрсету жұмыстың әдістемелік мазмұнын толықтыра түсер еді. 2. Болашақта модельдердің әртүрлі тақырыптық салалардағы мәтіндерді жіктеу қабілетін қосымша зерттеу ұсынылады. Аталған ескертулер диссертацияның құндылығын төмендетпейді және олар кеңес ретінде ұсынылған. Диссертациялық зерттеу жұмысы аяқталған зерттеу болып табылады.	
12.	Докторант мақалаларының зерттеу тақырыбы бойынша ғылыми деңгейі (диссертация мақалалар сериясы нысанында қорғалған жағдайда ресми рецензенттер докторанттың зерттеу тақырыбы бойынша әр мақаласының ғылыми деңгейін зерделейді)	Докторант мақалаларының зерттеу тақырыбы бойынша нәтижелері 10 ғылыми басылымда жарияланды. 1 мақала Scopus базасында индекстелетін басылымда, 3 мақала Scopus базасында индекстелетін конференцияларда, 2 мақала ҚР БҒМ Білім және ғылым саласындағы бақылау комитеті ұсынған журналдарында жарық көрді. 4 мақала халықаралық конференциялар материалдарында жарияланып, ҚР авторлық құқықпен қорғалатын объектілерге құқықтарды мемлекеттік тізілімге мәліметтерді енгізу туралы 1 куәлік алынды. Жарияланымдар халықаралық рецензенттердің сараптамалық бағалауынан өтті, бұл жүргізілген зерттеудің ғылыми жаңалалығы мен практикалық маңыздылығын растайды.	
13.	Ресми рецензенттің шешімі (осы Үлгі ереженің 28-тармағына сәйкес)	Қорытындылай келе, Сағынай Мөлдірдің «8D06301 – Ақпараттық қауіпсіздік жүйелері» білім беру бағдарламасы аясында орындалған «Қазақ тіліндегі деструктивті веб-контентті классификациялауға арналған тереңдетіп оқыту үлгілерін жасау» тақырыбындағы диссертациялық жұмысы қойылған мақсаттары, міндеттері, көлемі, ғылыми жаңалығы және тәжірибелік маңыздылығы тұрғысынан диссертациялық еңбектерге қойылатын талаптарға толық сәйкес келеді. Жұмыс толыққанды ғылыми зерттеу болып сасаналады. Сондықтан докторант Сағынай Мөлдір «8D06301 – Ақпараттық қауіпсіздік жүйелері» білім беру бағдарламасы бойынша философия докторы (PhD) дәрежесін алуға лайықты деп есептеймін.	

Ресми рецензент:
Qazaq AI Research University, PhD, профессор **Жукабаева Т.К.**