

**«6D070300 – Ақпараттық жүйелер» мамандығы бойынша философия докторы (PhD) дәрежесіне іздену үшін ұсынылған Кәрібаева Айдана Сейілғазықызының «Нейронды машиналық аударма үшін қазақ тіліндегі мәтіндерді морфологиялық сегментациясының модельдері мен әдістерін құру және зерттеу» тақырыбындағы диссертациялық жұмысына ресми рецензенттің**

**СЫН-ПІКІРІ**

р/н №	Критерийлер	Критерийлер сәйкестігі	Ресми рецензенттің ұстанымы
1.	Диссертация тақырыбының (бекіту күніне) ғылымның даму бағыттарына және/немесе мемлекеттік бағдарламаларға сәйкес болуы	<p>1.1 Ғылымның даму бағыттарына және/немесе мемлекеттік бағдарламаларға сәйкестігі:</p> <p>1) Диссертация мемлекет бюджетінен қаржыландырылатын жобаның немесе нысаналы бағдарламаның аясында орындалған (жобаның немесе бағдарламаның атауы мен нөмірі);</p> <p>2) Диссертация басқа мемлекеттік бағдарлама аясында орындалған (бағдарламаның атауы)</p> <p>3) Диссертация Қазақстан Республикасының Үкіметі жанындағы Жоғары ғылыми-техникалық комиссия бекіткен ғылым дамуының басым бағытына сәйкес (бағытын көрсету)</p>	<p>«Нейронды машиналық аударма үшін қазақ тіліндегі мәтіндерді морфологиялық сегментациясының модельдері мен әдістерін құру және зерттеу» тақырыбындағы диссертация ғылымның даму бағытына сәйкес. Қазіргі таңда кез келген нейронды машиналық аудармада мәтінге дұрыс және тиімді сегментация жасау өзекті мәселелердің бірі болып табылады.</p> <p>Диссертация Қазақстан Республикасы Білім және Ғылым Министрлігінің АР05131415 «Қазақ тілінің нейронды машиналық аудармасын құру және зерттеу» (2017-2020 жж., ғылыми жетекші: Тукеев У.А.) және АР08052421 «Машиналық аудармада Қазақ тілінің пост-редакциялау жүйесін зерттеу және әзірлеу» (2020-2022 жж., ғылыми жетекші: Рахимова Д.Р.) тақырыптары бойынша мемлекет бюджетінен қаржыландырылған жобалары аясында жүргізілген.</p>
2.	Ғылымға маңыздылығы	Жұмыс ғылымға елеулі үлесін қосады/қоспайды, ал оның маңыздылығы <u>ашылған</u> /ашылмаған.	Диссертацияда қазақ тілін компьютерлік өңдеу зерттеу бағытына және қазақ тілінің нейронды машиналық аударма облысына үлес қосады. Диссертацияны орындау барысында алынған нәтижелер ғылыми тұрғыдан маңызды. Зерттеудің маңыздылығы толығымен ашылған.
3.	Өзі жазу	Өзі жазу деңгейі:	Диссертацияның мазмұны салыстырмалы түрде жоғары деп

	принципі	<ul style="list-style-type: none"> <li>1) жоғары;</li> <li>2) орташа;</li> <li>3) төмен;</li> <li>4) өзі жазбаған</li> </ul>	бағалауға мүмкіндік береді. Диссертацияны жазуы, сипаттауы, рәсімдеуі жоғары деңгейде көрсетілген.
4.	Ішкі бірлік принципі	<p>4.1 Диссертация өзектілігінің негіздемесі:</p> <ul style="list-style-type: none"> <li>1) негізделген;</li> <li>2) жартылай негізделген;</li> <li>3) негізделмеген.</li> </ul>	Диссертация актуалды тақырыпқа орындалған және өзектілігі толығымен негізделген. Қазіргі таңда қазақ тілінің нейронды машиналық аудармасында сегментациялау жасаудың өзектілігі негізделген. Машиналық аударма мәселесі әлі жеткілікті жоғары деңгейде кәсіби аудармаға жақын шешілмегендіктен, машиналық аударма мәселесі өте өзекті болып табылады. Морфологиялық сегментациялау нейронды машиналық аудармада оқыту кезінде туындайтын жады мәселесін шешеді және қазақ тіліндегі мәтінді дұрыс сегменттерге бөлінуінің тиімділін арттырады, бұл диссертациялық жұмыстың машиналық аударма мәселесін шешу, табиғи тілді түсіну сияқты жасанды интеллекттің басқа да өте маңызды мәселелерін шешуге жол апаратындығын атап өткен жөн.
		<p>4.2 Диссертация мазмұны диссертация тақырыбын айқындайды</p> <ul style="list-style-type: none"> <li>1) айқындайды;</li> <li>2) жартылай айқындайды;</li> <li>3) айқындамайды</li> </ul>	Диссертацияның құрылымы толығымен тақырыпты айқындайды. Диссертацияның барлық бөлімдері зерттеу тақырыбының мазмұнын айқындайды.
		<p>4.3. Мақсаты мен міндеттері диссертация тақырыбына сәйкес келеді:</p> <ul style="list-style-type: none"> <li>1) сәйкес келеді;</li> <li>2) жартылай сәйкес келеді;</li> <li>3) сәйкес келмейді</li> </ul>	Зерттеудің мақсаты және міндеттері диссертация тақырыбына сәйкес келеді. Зерттеудің мақсаты: лингвистикалық ерекшеліктері негізінде қазақ тілінің нейронды машиналық аудармасының сапасын жақсарту үшін модельдер, алгоритмдер мен бағдарламалық жасақтама құру болып табылады. Жұмыстың міндеттері толығымен тақырыпқа сәйкес.
		<p>4.4. Диссертацияның барлық бөлімдері мен құрылысы логикалық байланысқан:</p> <ul style="list-style-type: none"> <li>1) толық байланысқан;</li> <li>2) жартылай байланысқан;</li> <li>3) байланыс жоқ</li> </ul>	Диссертацияның барлық бөлімдері, қорғауға шығарылған негіздеме бір-бірімен логикалы түрде байланысқан. Зерттеу тұжырымға сәйкес бөлімдерден тұрады. Диссертацияның құрылымы бұл дейгейдегі біліктілік жұмысына қойылатын талаптарға сәйкес
		<p>4.5 Автор ұсынған жаңа шешімдер (қағидаттар, әдістер) дәлелденіп,</p>	Автормен жұмыста алғаш рет CSE-моделіне негізделген морфологиялық сегментациялаудың жаңа шешімі ұсынылған.

		<p>бұрыннан белгілі шешімдермен салыстырылып бағаланған:</p> <p>1) сыни талдау бар;</p> <p>2) талдау жартылай жүргізілген;</p> <p>3) талдау өз пікірін емес, басқа авторлардың сілтемелеріне негізделген</p>	
5.	Ғылыми жаңашылдық принципі	<p>5.1 Ғылыми нәтижелер мен қағидаттар жаңа болып табыла ма?</p> <p>1) толығымен жаңа;</p> <p>2) жартылай жаңа (25-75% жаңа болып табылады);</p> <p>3) жаңа емес (25% кем жаңа болып табылады)</p>	Алынған ғылыми нәтижелері жаңа болып табылады. Нәтижелер алғаш қазақ тілі морфологиясының есептеуіш CSE (Complete Set of Endings) - моделі негізінде алынған. Зерттеу барысында алынған нәтижелер мен қағидаттар жаңа.
		<p>5.2 Диссертацияның қорытындылары жаңа болып табыла ма?</p> <p>1) толығымен жаңа;</p> <p>2) жартылай жаңа (25-75% жаңа болып табылады);</p> <p>3) жаңа емес (25% кем жаңа болып табылады)</p>	Диссертацияның қорытындылары мен қорғауға шығарылған тұжырымдар жаңа болып табылады. Бұл сонымен қатар автор рецензияланған журналдарға шығарған баспа жұмыстарымен дәлелденеді. Жасалған зерттеу жұмысының нәтижелері толығымен жаңа.
		<p>5.3 Техникалық, технологиялық, экономикалық немесе басқару шешімдері жаңа және негізделген бе?</p> <p>1) толығымен жаңа;</p> <p>2) жартылай жаңа (25-75% жаңа болып табылады);</p> <p>3) жаңа емес (25% кем жаңа болып табылады)</p>	Қойылған есепті шешуде қол жеткізу үшін қолданылған техникалық шешімдер жаңа болып табылады.
6.	Негізгі қорытындылардың негізділігі	Барлық қорытындылар ғылыми тұрғыдан қарағанда ауқымды дәлелдемелерде негізделген/негізделмеген (qualitative research және өнертану және гуманитарлық бағыттары бойынша)	Зерттеу жұмысында алынған барлық нәтижелердің қорытындылары ғылыми тұрғыдан негізделген.
7.	Қорғауға	Әр қағидат бойынша келесі сұрақтарға	7.1 Диссертацияны қорғауға шығарылатын негізгі қағидаттары

	шығарылған негізгі қағидаттар	жауап беру қажет: 7.1 Қағидат дәлелденді ме? <u>1) дәлелденді;</u> 2) шамамен дәлелденді; 3) шамамен дәлелденбеді; 4) дәлелденбеді 7.2 Тривиалды ма? 1) ия; <u>2) жоқ</u> 7.3 Жаңа ма? <u>1) ия;</u> 2) жоқ 7.4 Қолдану деңгейі: 1) тар; 2) орташа; <u>3) кең</u> 7.5 Мақалада дәлелденген бе? <u>1) ия;</u> 2) жоқ	дәлелденген. Диссертацияда тривиальдылық элементтері жоқ (7.2). Қорғауға шығарылған негізгі тұжырымдар жаңа (7.3). Қолданудың кең (7.4) деңгейіне ие, құрылған модельді негізге ала отырып, басқа түркі тілдер тобындағы тілдерге стемминг, морфологиялық сегментация, морфологиялық талдау есептерінде де қолдануға болады. Зерттеудің нәтижелері 21 мақалада жарық көрген (7.5). Оның 9-ы Web of Science және Scopus базасында шыққан.
8.	Дәйектілік принципі Дереккөздер мен ұсынылған ақпараттың дәйектілігі	8.1 Әдістеменің таңдауы - негізделген немесе әдіснама нақты жазылған <u>1) ия;</u> 2) жоқ 8.2 Диссертация жұмысының нәтижелері компьютерлік технологияларды қолдану арқылы ғылыми зерттеулердің қазіргі заманғы әдістері мен деректерді өңдеу және интерпретациялау әдістемелерін пайдалана отырып алынған: <u>1) ия;</u> 2) жоқ 8.3 Теориялық қорытындылар, модельдер, анықталған өзара байланыстар және заңдылықтар эксперименттік зерттеулермен дәлелденген және	Диссертацияда қолданылған әдіснаманың таңдауы жақсы негізделген және диссертацияның сәйкес бөлімінде сипатталған. Диссертацияның нәтижелері ғылыми зерттеудің заманауи әдістерін және компьютерлік технологияларды қолдана отырып деректерді өңдеу және интерпретациялау әдістерін қолдану арқылы алынған. Зерттеуде теориялық тұжырымдар мен қорытындылар, құрылған модельдер мен олардың өзара байланыстары эксперименттік зерттеулермен дәлелденген және расталған. Сонымен қатар, қорытындылар, модельдер ізденушімен құрылған компьютерлік

		расталған (педагогикалық ғылымдар бойынша даярлау бағыттары үшін нәтижелер педагогикалық эксперимент негізінде дәлелденеді): 1) ия; 2) жоқ	програамалық жасақтамамен де расталған.
		8.4 Маңызды мәлімдемелер нақты және сенімді ғылыми әдебиеттерге сілтемелермен <u>расталған</u> / ішінара расталған / расталмаған	Ізденушінің маңызды мәлімдемелері актуалды шынайы ғылыми әдебиеттерге сілтемелермен расталған. Жұмыста көрсетілген пайдаланған әдебиеттер тізімі зерттеу саласына сәйкес келеді.
		8.5 Пайдаланылған әдебиеттер тізімі әдеби шолуға <u>жеткілікті</u> /жеткіліксіз	Диссертацияда пайдаланған әдебиеттер саны әдеби шолу үшін жеткілікті болып табылады. Диссертацияның шолу бөлімінде дүние жүзі бойынша сегментацияға қатысты жұмыстарға, мақалаларға толыққанды әдебиеттік шолу жасалынған.
9	Практикалық құндылық принципі	9.1 Диссертацияның теориялық маңызы бар: 1) ия; 2) жоқ	Диссертацияның теориялық маңызы бар. Диссертация табиғи тілдерді өндеуде, қазақ тілін компьютерлік өндеуде теориялық жағынан маңызды.
		9.2 Диссертацияның практикалық маңызы бар және алынған нәтижелерді практикада қолдану мүмкіндігі жоғары: 1) ия; 2) жоқ	Диссертацияның практикалық маңызы бар. Құрылған морфологиялық сегментацияны қазақ тілі бір бағыт ретінде бар нейронды машиналық аудармада алдын-ала өңдеудің сатысы ретінде қолданыла алынады, бұл аударма жүйесінің сапасын арттыруға көмектеседі.
		9.3 Практикалық ұсыныстар жаңа болып табылады? 1) толығымен жаңа; 2) жартылай жаңа (25-75% жаңа болып табылады); 3) жаңа емес (25% кем жаңа болып табылады)	Зерттеуде қазақ тілінің лингвистикалық ерекшеліктерін ескере отырып, CSE-моделі жетілдірілді. Аталған модель негізінде ұсынылған морфологиялық сегментациялаудың әдісін қазақ тілінің нейронды машиналық аудармасында тәжірибелер жүргізу арқылы практикалық қолданысы көрсетілген. Диссертацияда көрсетілген практикалық ұсыныстар толығымен жаңа болып табылады.
10.	Жазу және ресімдеу сапасы	Академиялық жазу сапасы: 1) жоғары; 2) орташа; 3) орташадан төмен;	Диссертация өте жоғары зерттеу деңгейінде жазылған, техникалық стильге ие. Диссертацияның жалпы ресімделуі философия докторы (phD) дәрежесін алу үшін докторлық диссертацияларға қойылатын талаптарға сәйкес келеді

Ескертулер мен ұсыныстар:

Диссертацияда қазақ тілінің CSE-моделі құрылды, бірақ кей жерлерде ғылыми терминдердің қазақша аудармасын нақтылау керек (юниграмм, униграмм). Осы тақырып бойынша болашақ зерттеуді жалғастыру үшін жетілдірілген CSE-моделін басқа да Түркі тілдері үшін қолданысын қарастыруы ұсынылады.

Қорытынды:

Кәрібаева Айдана Сейілғазықызының «Нейронды машиналық аударма үшін қазақ тіліндегі мәтіндерді морфологиялық сегментациясының модельдері мен әдістерін құру және зерттеу» тақырыбындағы диссертациясы «Дәрежелерді беру қағидаларын бекіту» талаптарына сай келеді, ал оның авторы «6D070300 – Ақпараттық жүйелер» мамандығы бойынша философия докторы (pHd) дәрежесіне лайық.

Ресми рецензент:

Д. Серікбаев атындағы Шығыс Қазақстан техникалық университетінің  
қауымдастырылған профессоры,  
техника ғылымдарының кандидаты

Кумаргажанова С. К.

