

https://doi.org/10.51885/3134-8025_IICS_2026_2_8

XFTAP 28.23.00

МУЛЬТИМОДАЛЬДЫ СИГНАЛДАРДАҒЫ ЭМОЦИЯЛАРДЫҢ УАҚЫТША ТҰРАҚТЫЛЫҒЫН БАҒАЛАУ

ОЦЕНКА ВРЕМЕННОЙ СТАБИЛЬНОСТИ ЭМОЦИЙ В МУЛЬТИМОДАЛЬНЫХ СИГНАЛАХ

EVALUATION OF TEMPORAL STABILITY OF EMOTIONS IN MULTIMODAL SIGNALS

С.Ә. Кудубаева ^{1,2}, Е. Сексенбаев ², Ә.Н. Дайырбаева ^{1,3*}, А.С. Еримбетова ¹,
Ж.А. Бурибаев ^{1,4}

¹ҚР ҒЖБМҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы қ., Қазақстан

²Л.Н. Гумилев атындағы Еуразиялық ұлттық университеті, Астана қ., Қазақстан

³Қ.И.Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы қ., Қазақстан

⁴Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

*Жауапты автор: Дайырбаева Эльмира Нурбекқызы, e-mail: nurbekkyzelmira@gmail.com

Түйінді сөздер:

Эмоцияны
мультимодальды тану,
уақытша тұрақтылық,
кросс-модальды зейін,
трансформаторлар,
уақытша консистенцияны
жүйелеу, Херст
көрсеткіші,
автокорреляция,
бұрыштық жылдамдық
(траектория), шартты
кездейсоқ өрістер(CRF),
модальділік-беріктіктің
болмауы.

ТҮЙІНДЕМЕ

Бұл мақалада мультимодальды мәліметтердегі эмоциялық күйлердің уақытша тұрақтылығын бағалауға арналған әдістемелік тәсіл ұсынылады. Басқа жұмыстардан айырмашылығы, біз көп деңгейлі формализацияны (сигнал → функция → жасырын кеңістік → болжау) және көрсеткіштер жиынтығын (класс ауыстырғыштарының саны, эпизодтардың ұзақтығы және Херст көрсеткіші) енгіземіз. Тұрақтылық «Плутчик» моделі негізінде түсіндірілетін полярлық кеңістіктегі траекториялар деңгейінде бағаланады. Жұмыста уақытша тұрақтылықты сипаттайтын көрсеткіштер жүйесі ұсынылып, олар бірнеше мультимодальды архитектура-ларда салыстырмалы түрде талданады. Бағалау хаттамасына динамиктерден тәуелсіз бөлімдер, стандартты көрсеткіштер (F1/UAR, CCC/MAE) және уақытша тұрақтылық көрсеткіштері енгізілді. Қосымша ретінде MELD деректер жиынтығының теңгерілген шағын ішкі жиынтығында text-only pilot-эксперимент жүргізілді. Ұсынылған бағалау тәсілі модельдердің уақыт бойынша тұрақтылық қасиеттерін дәлірек сипаттауға мүмкіндік береді. Практикалық құндылық адам мен компьютердің өзара әрекеттесу жүйелеріндегі және көмекші технологиялардағы мінез-құлықтың болжамдылығы мен түсіндірілуін арттыруда жатыр.

Ключевые слова:

Распознавание эмоций в
мультимодальных
данных, временная
стабильность, кросс-
модальное внимание,

АННОТАЦИЯ

В данной статье предлагается методический подход к оценке временной стабильности эмоциональных состояний в мультимодальных данных. В отличие от других работ, мы вводим многоуровневую формализацию (сигнал → функция → скрытое пространство → предсказание) и набор показателей, включающий



трансформеры,
систематизация
временной
консистентности,
показатель Херста,
автокорреляция, угловая
скорость (траектория),
условные случайные поля
(CRF), отсутствие
модальной устойчивости

число переключений классов, длительность эпизодов и показатель Херста. Стабильность оценивается на уровне траекторий в полярном пространстве, интерпретируемом на основе модели Плутчика. В работе предлагается система показателей, описывающих временную стабильность, и проводится их сравнительный анализ на нескольких мультимодальных архитектурах. Протокол оценки включает разбиения, независимые от дикторов, стандартные метрики (F1/UAR, CCC/MAE) и показатели временной стабильности. В качестве дополнительной проверки был проведён text-only pilot-эксперимент на сбалансированной небольшой подвыборке датасета MELD. Предложенный подход к оценке позволяет точнее характеризовать свойства временной устойчивости моделей. Практическая ценность работы заключается в повышении предсказуемости и интерпретируемости поведения систем человеко-компьютерного взаимодействия и вспомогательных технологий.

keywords:

Multimodal emotion
recognition, temporal
stability, cross-modal
attention, transformers,
systematization of temporal
consistency, Hurst
exponent, autocorrelation,
angular velocity
(trajectory), conditional
random fields (CRF), lack of
modality robustness.

ABSTRACT

This paper proposes a methodological approach for evaluating the temporal stability of emotional states in multimodal data. Unlike previous studies, we introduce a multi-level formalization (signal → function → latent space → prediction) and a set of indicators including the number of class switches, episode duration, and the Hurst exponent. Stability is assessed at the level of trajectories in a polar space interpreted on the basis of Plutchik's model. The study proposes a system of metrics describing temporal stability and provides a comparative analysis of these metrics across several multimodal architectures. The evaluation protocol includes speaker-independent splits, standard metrics (F1/UAR, CCC/MAE), and temporal stability indicators. As an additional validation step, a text-only pilot experiment was conducted on a balanced small subset of the MELD dataset. The proposed evaluation approach makes it possible to characterize the temporal stability properties of models more accurately. The practical value of the work lies in improving the predictability and interpretability of behavior in human-computer interaction systems and assistive technologies.

КІРІСПЕ

Эмоцияларды тану – адам мен машина арасындағы өзара әрекеттесуді табиғи әрі тиімді ету үшін маңызды зерттеу бағыты. Соңғы жылдары эмоцияны тану әдістері бейне, аудио және мәтін сияқты әртүрлі деректер көздерін қамтитын мультимодальды тәсілдерге негізделіп дамып келеді. Дегенмен, эмоциялар уақыт бойынша тұрақты құбылыс емес: олар бір сәтте өзгеріп, әртүрлі қарқындылықта көрінуі мүмкін екендігі көптеген зерттеулерде көрсетілген. Сондықтан эмоциялардың уақытша тұрақтылығын бағалау эмоцияны дұрыс модельдеу мен сенімді тануда шешуші рөл атқарады. Бұл мәселе эмоционалдық күйлердің динамикасын сипаттайтын көрсеткіштерді (мысалы, ауысулар саны, эпизод ұзақтығы, автокорреляция, Херст көрсеткіші) қолдануды талап етеді. Жоғарыда келтірілген тәсілдер тек тану дәлдігін арттырып қана қоймай, эмоциялардың табиғи ағымын ескеруге мүмкіндік береді, соның нәтижесінде мультимодальды эмоцияларды талдау жүйелері неғұрлым сенімді әрі түсінікті бола түседі. Мақала мазмұнын ашу барысында мультимодальды эмоцияларды талдау жүйелеріне негізделген зерттеулерге талдау жасадық.

Jayasinghe, Wong және Nugaliyadde (2026) өз кезегінде Speech Emotion Recognition (SER) – адамның сөйлеуінен эмоционалды күйлерді анықтау әдісін қарастырды. Негізінен,

Automatic SER (ASER) саласында эмоцияны болжауда қолданылатын машиналық оқыту (ML) тәсілдерінің интерпретабельділігі мен түсіндірілуіне назар аударылды. Авторлар ASER модельдерінің сенімділігі мен ашықтығы психикалық денсаулықты бақылау сияқты нақты қолданбалар үшін маңызды екенін көрсетті. Сол сияқты бағалау метрикаларын стандарттау, мультимодальды деректерді біріктіру және модельдердің тұрақтылығын арттыру мәселелері әлі толық шешілмегені де мақалада көрсетілген. Осылайша, түйіндейтін болсақ, интерпретабельділік пен түсіндіруді жетілдірудің ASER жүйелерін сенімді енгізу үшін шешуші маңызға ие екенін көрсетілген.

Қарастырып отырған келесі мақалада (Wang және басқ., 2026) мультимедиялық технологиялардың қарқынды дамуы адамдардың интернетте өз пікірін бейнелер арқылы жиі жеткізуіне ықпал еткені туралы айтылған. Осыған байланысты қазіргі таңда эмоцияларды визуалды талдау (VEA) барлық әлем зерттеушілерінің назарын аударған. Эмоциялар субъективті құбылыс болғандықтан, олар бейненің түсі, жарықтығы, бет әлпеті, объектілері мен көріністері сияқты түрлі атрибуттарға тәуелді болатындығын ескеріміз керек. Осы мәселені шешу үшін авторлар эмоцияға қатысты атрибуттық ақпаратты кешенді түрде алуға арналған атрибуттарға негізделген терең оқыту желісін (AttDNet) ұсынды. Ұсынылған желі бірнеше модульдерден тұратын атрибуттық тармақ пен бет әлпеті туралы ақпарат алатын тармақтан құралады. Атрибуттық тармақтың модульдері әртүрлі ақпаратты өңдейді, атап айтсақ CSCAM модулі бейненің түсі мен негізгі объектілерін, LEconv модулі объектілер контекстін, ал MLPooling модулі көрініс туралы мәліметтерді шығарады. Сонымен қатар, бір эмоциялық категорияға жататын үлгілерді жақындату үшін CenterLoss әдісін қолданған.

Zhang және басқ. (2026) мақаласында мультимодальды сарказмды анықтау мәселесін қарастырды. Аталған бағыттағы негізгі қиындық – әртүрлі модальдар арасындағы өзара әрекеттесуді толық ескеру қажеттілігі болып табылады. Авторлар сарказмды тануда тақырыптардың маңызды рөл атқаратынын негізге ала отырып, тақырыпқа негізделген енгізу мүмкіндіктерін жетілдіру әдісін ұсынды. Ұсынылған әдіс екі деректер жиынтығында (бейне негізіндегі және твиттер негізіндегі) және он екі негізгі мультимодальды сарказмды анықтау моделінде сынақтан өткізілген. Эксперименттік нәтижелер барлық базалық модельдерде көрсеткіштердің едәуір жақсарғанын көрсетті. Осылайша, авторлар ұсынған тәсілдің мультимодальды сарказмды анықтауда тиімділігін дәлелдеген.

Pradhan және Srivastava (2026) мультимодальды физиологиялық деректер негізінде эмоцияларды тануға арналған жаңа модельді ұсынған. Алдымен деректер арнайы сүзгілеу әдістері арқылы тазартылып, кейін SCA-GAN генеративті желісі арқылы аугментация жүргізілген. Модель кеңістікке бағытталған назар (SA) мен арнаға бағытталған назарды (CA) қолданып, деректерді теңгерімдеуді және ерекшеліктерді толық қамтуды қамтамасыз етті. Жүргізілген зерттеудің соңғы этапында эмоциялар HPC-CNN гибридті нейрондық желісі арқылы жіктелген. Зерттеу барысында ашық WESAD деректер жиынтығы пайдаланылды. Нәтижелер ұсынылған әдістің жоғары тиімділігін көрсетті: дәлдік – 99,4 %, нақтылық – 98,04 %, ерекшелік – 99,56 %, F1 көрсеткіші – 98,44 %, ал сезімталдық – 98,84 % болды. Қорытындылай келе, (Pradhan және Srivastava, 2026) ұсынған тәсіл мультимодальды физиологиялық сигналдар арқылы эмоцияларды сенімді әрі дәл анықтауға мүмкіндік беретінін дәлелдеді.

Қарастырылған (Qi, Huang және Zhang, 2022) жұмыста сөйлеу эмоцияларын тануға арналған әдіс ретінде CapsNet капсульді нейрондық желісі қолданылған. Модельдің тиімділігін арттыру үшін сөйлеу деректеріне «ақ» шу қосу және дауыс өзгерту арқылы аугментация жүргізілген. Жүйеге енгізілген параметрлерге MEL-спектр бейнелерімен қатар дыбыс көзі, дауыс жолы сияқты ерекшеліктер енгізілген. Эксперимент нәтижелері көрсеткендей, неміс тіліндегі EMODB корпусында төрт эмоцияны (бейтарап, зеріккен, ашу,

қуаныш) тану дәлдігі 99,69%, ал вьетнам тіліндегі ВКЕмо корпусында (бейтарап, мұң, ашу, қуаныш) 94,23% болды. Ең жоғары көрсеткіштер MEL-спектр бейнелерін негізгі жиілікке қатысты сипаттамалармен біріктіргенде алынды.

Van, Nguyen және Le (2022) жұмысында эмоцияларды тануға арналған капсулалық нейрондық желі негізіндегі тәсіл қарастырылған. Авторлар модельдің эмоциялық белгілерді ықшам әрі тиімді ұсыну мүмкіндігін көрсетіп, оның дәстүрлі архитектуралармен салыстырғанда перспективалы екендігін негіздейді.

Авторлар (Liu және басқ. 2023) сөйлеу эмоцияларын танудағы деректердің әртүрлі адамдар арасында тең бөлінбеу мәселесін шешу үшін теңгерімді аугментацияланған үлгілерді қолданды. Бұл тәсіл үш арналы LOG-MEL спектрограммаларын негізге алып, эмоциялық категориялар арасындағы үлгі теңгерімсіздігін азайтып, нейрондық желіге жеткілікті деректер ұсынуға мүмкіндік берді. Ерекшеліктерді әртараптандыру үшін уақыттық және жиіліктік сүзгілер қолданылды. Модельде белгілерді алу үшін CNN мен назарға негізделген екі бағытты ұзақ-қысқа мерзімді жады желісі (ABLSTM) біріктірілді, ал өнімділікті арттыру үшін көптапсырмалық оқыту енгізілді. Жеті көмекші тапсырма қарастырылып, тәжірибелік салыстыру арқылы ең тиімділері таңдалды. Ұсынылған әдіс IEMOCAP және MSP-IMPROV деректер жиынтықтарында тексерілді. Нәтижесінде IEMOCAP базасында WAR=70,27 % және UAR=66,27 %, ал MSP-IMPROV базасында WAR=60,90 % және UAR=61,83 % көрсеткіштері алынған және нәтижелер әдістің алдыңғы тәсілдерге қарағанда тиімділігін байқатқан.

Sun және басқ. (2019) терең нейрондық желілердің (DNN) дамуына сүйене отырып, сөйлеу эмоцияларын танудың DNN-decision tree SVM моделі негізіндегі әдісін ұсынған. Қолданылған тәсіл сөйлеу сигналдарынан терең эмоциялық ақпаратты алуға және бір-бірімен шатастырылатын эмоциялардан айқынырақ ерекшеліктерді бөлуге мүмкіндік береді. Біріншіден эмоциялардың шатасу дәрежесі есептеліп, соның негізінде decision tree SVM құрылымы жасалды. Кейін әр түрлі эмоциялар тобы үшін бөлек DNN модельдері оқытылып, алынған «bottleneck» ерекшеліктер decision tree-дегі әрбір SVM-ді үйретуге қолданылды. Авторлар ұсынып отырған әдіс Қытай ғылым академиясының эмоциялық корпусында бағаланды. Эксперимент нәтижелері ұсынылған модельдің эмоцияларды тану дәлдігі дәстүрлі SVM-нен 6,25%, ал DNN-SVM әдісінен 2,91% жоғары болғанын көрсетті. Бұл тәсіл эмоциялар арасындағы шатасуды азайтып, сөйлеу эмоцияларын танудың нақтылығын арттыратынын көрсетілді.

Қарастырылған жұмыстар эмоцияларды тануда мультимодальды біріктірудің, назар механизмдерінің және трансформаторлық архитектуралардың тиімділігін көрсетеді. Алайда бұл зерттеулердің басым бөлігі стандартты дәлдік метрикаларына сүйенеді және эмоция болжамдарының уақыт бойынша тұрақтылығын жеке зерттеу нысаны ретінде қарастырмайды. Сонымен қатар, уақытша тұрақтылықты бағалау үшін қолданылатын көрсеткіштер жүйесі біріздендірілмеген. Осы олқылық аталған жұмыстың негізгі зерттеу уәжін құрайды.

Біздің мақала тақырыбы бағытындағы зерттеулер осы уақытқа дейін негізінен дәйекті өңдеу үшін қайталанатын нейрондық желілерді пайдаланды. Жалпы қондырғы ResNet-50-ден визуалды мүмкіндіктерді көрсетті (He және басқ., 2016) және 1D-CNN-ден аудио мүмкіндіктер (Kiranyaz және басқ., 2019), оларды біріктіріп, нәтижені LSTM-ге жіберіп, әр сегмент үшін эмоцияны болжады. Осыған ұқсас құрылғылар әсерді үздіксіз бағалау үшін пайдаланылды, яғни CNN-нің алдын ала дайындалған бет-әлпеті мен конволюциялық аудио модельдері мүмкіндіктерді қамтамасыз етті, ал бағалау көрсеткіштері уақытша ақпаратты модельдеу өнімділікті арттыратынын дәйекті түрде көрсетті.

Алайда, бұл модельдер көбінесе кросс-модальды өзара әрекеттесуді тиімсіз басқарды. Сөйлесу эмоциясын тану үшін графикке негізделген әдістер енгізілді.

DialogueGCN (Ghosal және басқ., 2019) сөйлеушілерді графикалық түйіндер ретінде және олардың өзара әрекеттесуін жиектер ретінде ұсынады, айтылымдар бойынша аффект ағынын модельдейді. DialogueRNN (Majumder және басқ., 2019) әр қатысушының жағдайын бақылау үшін жеке қайталанатын модульдерді қосады; үш GRU тізбегімен ол кімнің сөйлейтінін және қандай контексте сөйлейтінін нақты модельдейді, бұл жіктеу дәлдігін жақсартты.

Зейінге негізделген модельдер ақпараттық уақытша сегменттер мен модальдарға назар аударуға алады. Кросс-модальды зейінді қолданатын әдістер әсіресе тиімді бола бастады. Мұндай тәсілдер әртүрлі модальдар арасындағы өзара байланысты айқындап, біріктірілген көріністің сапасын арттыруға мүмкіндік береді. Ұзақ қашықтыққа тәуелділікті түсіруге қабілетті трансформаторлар өрісті одан әрі жетілдірді. Tsai және басқ., (2019) жұмысында ұсынылған «мультимодальды трансформатор» (MulT) бағытталған, жұптық кросс-модальды зейінді енгізді: әрбір ағын (мысалы, аудио – бейне, мәтін – бейне) назар аудару блоктары арқылы көріністермен алмасады, бұл асинхронды реттіліктерді нақты уақыт режимінде бекітпестен туралауға мүмкіндік береді. Li және басқ., (2022) MulT CMU-MOSEI-де айтарлықтай жетістіктерге жетіп, эмоцияны болжау сапасын едәуір жақсарғанын атап өтті. Негізгі қасиет – оның әртүрлі реттілік ұзындықтары мен уақытша ауысуларға бейімделу қабілеті: әр түрлі уақыт кезеңдеріндегі кросс-модальды зейін аудио мен бейне дұрыс анықталмаған кезде де ортақ сигналдарды ашады. Тағы бір мысал: аудио – мәтінге арналған MMT (Мультимодальды Трансформатор); ол сөйлеуді (Гюберт) және мәтінді (BERT) ендіруді біріктіру үшін екі бағытты кросс-модальды зейінді қолданады, бір модальды екіншісімен қайта өлшейді (Moorthy және Moon, 2025). IEMOCAP-та MMT өлшенген дәлдікті 84,5 % дейін арттырды, бұл алдыңғы тәсілдерден асып түсті (Nakim, Marsland және Guesgen, 2013).

Иерархияны басып алу әрекеттері иерархиялық мультимодальды трансформаторлардың пайда болуына әкелді. Иерархиялық архитектуралар әртүрлі уақыттық деңгейлердегі белгілерді біріктіруге мүмкіндік береді. Мұндай тәсілдер қысқа мерзімді реакциялар мен ұзаққа созылатын эмоциялық үрдістерді қатар сипаттауға ыңғайлы. Ал Плутчик моделі (Plutchik, 1980) бұл жұмыста эмоциялық күйлерді интерпретациялауға арналған теориялық негіз ретінде қарастырылады. Бұл екі сатылы иерархия модальділікке тән мазмұнды және олардың корреляциясын сақтай отырып, біріктіру алдында әрбір модальділікті күшейтеді. AffWild2 және AFEW-VA бойынша жүргізілген бағалаулар HMATN бұрынғы базалық көрсеткіштерден асып түсетінін және шу мен жетіспейтін әдістерге төзімді екенін көрсетеді. Дегенмен, көп динамикті сөйлесу параметрлерінде графикалық және RNN негізіндегі тәсілдер (DialogueGCN, DialogueRNN және опциялар) динамиктер арасындағы өзара әрекеттесу мен аффективті динамиканы нақты модельдеу үшін өзекті болып қала береді. Осылайша, әдебиеттерді талдау мультимодальды эмоцияларды тану саласында дәлдік көрсеткіштері кең қолданылғанымен, эмоция болжамдарының уақыт бойынша тұрақтылығы жүйелі түрде жеткілікті зерттелмегенін көрсетеді. Осыған байланысты мақаланың ғылыми сұрағы келесідей тұжырымдалады: мультимодальды эмоцияларды тану модельдерінің қай түрі дәлдікпен қатар уақытша тұрақтылықты неғұрлым тиімді қамтамасыз етеді? Зерттеу гипотезасы бойынша, кросс-модальды және иерархиялық трансформаторлық архитектуралар ұзақ және қысқа мерзімді тәуелділіктерді жақсырақ модельдеу есебінен базалық RNN/GCN модельдеріне қарағанда уақытша тұрақтылық көрсеткіштері бойынша жоғары нәтиже береді.

ЗЕРТТЕУ МАТЕРИАЛДАРЫ МЕН ӘДІСТЕРІ

Уақытша тұрақтылықты бағалау үшін эмоция болжамдарының уақыт қатары бірнеше көрсеткіш арқылы сипатталды. Бірінші көрсеткіш — N_{switch} , яғни көршілес

уақыт нүктелеріндегі эмоция класының ауысу саны. Бұл мән неғұрлым төмен болған сайын, модель болжамы соғұрлым тұрақты деп қарастырылады. Екінші көрсеткіш — D_{avg} , бір эмоциялық күйдің орташа ұзақтығы; оның жоғары мәні эмоция эпизодтарының ұзағырақ сақталуын сипаттайды. Үшінші көрсеткіш — $AC(1)$ автокорреляциясы, ол уақыт бойынша көршілес күйлердің өзара ұқсастық дәрежесін бағалайды. Төртінші көрсеткіш — Херст көрсеткіші H , ол уақыт қатарындағы ұзақ мерзімді тәуелділікті сипаттайды. Егер $H > 0.5$ болса, қатарда персистенттілік байқалады; $H \approx 0.5$ болса, процесс кездейсоқ сипатқа жақын; ал $H < 0.5$ болса, анти-персистенттілік орын алады.

Эмоциялар контекстінде H көрсеткішінің жоғары мәні эмоционалдық күйдің өз бағытын сақтауға бейімділігін, ал төмен мәні жиі ауытқуларды білдіреді. Осы жұмыста уақытша тұрақтылықты сипаттау үшін N_{switch} , D_{avg} және $AC(1)$ көрсеткіштері негізгі талдау құралы ретінде қарастырылды. Херст көрсеткіші H теориялық тұрғыдан ұзақ мерзімді тәуелділікті сипаттайтын қосымша метрика ретінде енгізілді, алайда қысқа диалогтық тізбектер үшін оның тұрақтылығы шектеулі болуы мүмкін. Жалпы алғанда, аталған көрсеткіштер эмоция болжамдарының уақыт бойынша дәйектілігін жан-жақты бағалауға арналған бірыңғай талдау шеңберін құрайды. 1 – кестеде эмоциялар туралы мәліметтер жиынтығы және қимылдар келтірілген. Қосымша тәжірибелік тексеру ретінде MELD деректер жиынтығының теңгерілген шағын ішкі жиынтығында pilot-эксперимент жүргізілді. Ішкі жиынтық ресми train/dev/test бөлінісін сақтай отырып құрылды және әр эмоция класы бойынша теңгерілді. Мәтіндік белгі ретінде TF-IDF униграмма және биграмма сипаттамалары пайдаланылды. Базалық модельдер ретінде logistic regression, linear SVM және MLP қарастырылды. Әр тәжірибе бес random seed бойынша (13, 17, 23, 42, 77) қайталанды. Бағалау үшін macro-UAR, macro-F1, сондай-ақ уақытша тұрақтылық көрсеткіштері N_{switch} , D_{avg} және $AC(1)$ қолданылды. Нәтижелер үшін 95% confidence interval есептелді, ал baseline ретінде logistic regression алынып, paired significance test жүргізілді. Тұрақтылыққа қосымша тексеру ретінде мәтінге кездейсоқ бұзылу енгізілді (text corruption: 20 %, 40 % және 60 % сөздердің түсірілуі).

1-кесте. Эмоциялар туралы мәліметтер жиынтығы
және қимылдар ұсынылған

Деректер	Жиынының сипаттамасы және өлшем	Модальділіктері	Эмоция белгілері
IEMOCAP (Busso және басқ., 2008)	~12 сағаттық диалогтық диалогтар, ~5к сегменттер. Диалог контексті қолжетімді.	Аудио, бейне, (мимика, дене қалпы), мәтін.	9 санат (қуаныш*, қайғы, ашу, қорқыныш, таңдану, жиіркеніш, бейтараптық, + толқу, көңілсіздік); сонымен қатар әр сегмент үшін валенттілік, ерекше сезімге бөлену және үстемдік рейтингтері (1-5).
Aff-Wild2 (Kollias және басқ., 2019)	~50 сағат "жабайы табиғатта" (564 түрлі веб-бейнелер, 2,8 М кадрлар).	Аудио, бейне	Үздіксіз аннотация: бір кадрға [-1,1] диапазонындағы валенттілік пен ерекше сезімге бөлену. Категориялық: 7 негізгі эмоциялар.
MELD (Poria және басқ., 2019)	"Достар" телешоуынан 1433 көп динамикті диалог, барлығы 13 мың сөз. Диалогтың орташа ұзақтығы ~8-9 айналым, 4 спикерге дейін.	Аудио, бейне, мәтін	Бір сөзге эмоциялардың 7 санаты (ашу, жиіркеніш, қорқыныш, қуаныш, бейтараптық, қайғы, таңдану) + көңіл-күй (оң / бейтарап / теріс).

1-кестенің соңы

CMU-MOSEI (Liang, Zadeh және Morency, 2018)	23 453 бейне сегменті (youtube talks), әрқайсысы ~7 секунд, 1000 бірегей динамик; барлығы ~66 сағат.	Аудио, бейне, мәтін	Бірнеше белгілер: әр сегмент үшін қарқындылығы 0-3 болатын 6 эмоцияның (қуаныш, қайғы, ашу, қорқыныш, жиіркеніш, таңданыс) болуы/болмауы; + көңіл-күй [-3...+3]. (Эмоциялар бір-бірін жоққа шығармайды.)
Ескерту – авторлар құрастырған			



1-сурет. Ұсынылған бағалау сызбасы

Ескерту – авторлар алынған нәтижелер негізінде құрастырған

НӘТИЖЕЛЕР ЖӘНЕ ОЛАРДЫ ТАЛҚЫЛАУ

Бұл зерттеуде модельдердің үш тобы бағаланды: базалық RNN/GCN жүйелері, кросс-модальды назар аударатын трансформаторлар және иерархиялық трансформаторлар. Салыстыру IEMOCAP, Aff-Wild2, MELD және CMU-MOSEI деректер жиынтықтарында жүргізілді. Категориялық тапсырмалар үшін UAR, ал үздіксіз бағалау үшін CCC және MAE көрсеткіштері пайдаланылды. Негізгі салыстырмалы нәтижелер 2-кестеде ұсынылған.

2-кесте. Дәлдік (экспортталатын таңдау)

Датасет	Модельдер тобы	UAR	CCC	MAE
IEMOCAP	Baseline (RNN/GCN)	0,644	0,555	0,311
IEMOCAP	Transformer (Cross-Modal)	0,748	0,632	0,259
IEMOCAP	Hierarchical Transformer	0,726	0,594	0,253
Aff-Wild2	Baseline (RNN/GCN)	0,682	0,48	0,25
Aff-Wild2	Transformer (Cross-Modal)	0,666	0,646	0,243
Aff-Wild2	Hierarchical Transformer	0,718	0,714	0,249
MELD	Baseline (RNN/GCN)	0,619	0,484	0,226
MELD	Transformer (Cross-Modal)	0,721	0,629	0,277
MELD	Hierarchical Transformer	0,72	0,614	0,179
CMU-MOSEI	Baseline (RNN/GCN)	0,648	0,441	0,307
CMU-MOSEI	Transformer (Cross-Modal)	0,735	0,62	0,211
CMU-MOSEI	Hierarchical Transformer	0,754	0,597	0,189
Ескерту – авторлар құрастырған				

Дәлдіктен басқа, біз шынайы өзгерістерге сезімталдықты төмендетпей, болжамдардың тұрақтылығын сандық түрде анықтаймыз. 3-кесте деректері бойынша иерархиялық және кросс-модальды трансформаторлар базалық RNN/GCN модельдерімен салыстырғанда N_switch мәнін төмендетіп, H көрсеткішін арттырды. Бұл олардың эмоция болжамдарының уақыт бойынша неғұрлым тұрақты екенін көрсетеді. Полярлық эмоциялар кеңістігінде, яғни салыстырмалы жауап беру қабілетін ескергенде, «қатты» өзгерістер саны азаяды.

3-кесте. Уақытша тұрақтылық

Датасет	Модель	N_switch (↓)	H (0–1)
IEMOCAP	Baseline (RNN/GCN)	51,7	0,551
IEMOCAP	Transformer (Cross-Modal)	36	0,692
IEMOCAP	Hierarchical Transformer	20,4	0,744
Aff-Wild2	Baseline (RNN/GCN)	53,9	0,634
Aff-Wild2	Transformer (Cross-Modal)	24,6	0,682
Aff-Wild2	Hierarchical Transformer	22	0,709
MELD	Baseline (RNN/GCN)	48,5	0,611
MELD	Transformer (Cross-Modal)	27	0,744
MELD	Hierarchical Transformer	20,7	0,789
CMU-MOSEI	Baseline (RNN/GCN)	43,9	0,604
CMU-MOSEI	Transformer (Cross-Modal)	26,1	0,74
CMU-MOSEI	Hierarchical Transformer	34,5	0,74
Ескерту – авторлар құрастырған			

MELD деректер жиынтығының теңгерілген шағын ішкі жиынтығындағы қосымша pilot-эксперимент нәтижелері мәтіндік baseline-модельдердің мінез-құлқын нақты деректерде тексеруге мүмкіндік берді. Бұл тәжірибеде үш модель қарастырылды: logistic regression, linear SVM және MLP. Барлық модельдер бес random seed бойынша бағаланды. Орташа test-UAR мәндері тиісінше 0,263, 0,275 және 0,215 болды, ал test-F1 мәндері 0,263, 0,273 және 0,209 құрады. Осылайша, орташа сапа бойынша ең жақсы нәтижені linear SVM көрсетті, logistic regression оған жақын нәтиже берді, ал MLP ең төмен нәтижелерге ие болды.

4-кесте. MELD деректер жиынтығының теңгерілген шағын ішкі жиынтығындағы мәтіндік baseline-модельдердің нәтижелері

Модель	Test UAR	Test F1	N_switch	D_avg	AC(1)	p-value (UAR, logreg-пен салыстыру)
Logistic Regression	0,263	0,263	0,380	1,045	-0,456	–
Linear SVM	0,275	0,273	0,387	1,044	-0,468	0,108
MLP	0,215	0,209	0,382	1,046	-0,614	0,0067

Статистикалық тексеру linear SVM мен logistic regression арасындағы айырмашылықтың test-UAR ($p = 0,108$) және test-F1 ($p = 0,147$) бойынша статистикалық мәнді емес екенін көрсетті. Ал MLP моделі logistic regression-мен салыстырғанда test-UAR ($p = 0,0067$) және test-F1 ($p = 0,0076$) көрсеткіштері бойынша статистикалық мәнді түрде

төмен нәтиже берді. Уақытша тұрақтылық көрсеткіштері бойынша (N_{switch} , D_{avg} , $AC(1)$) модельдер арасында айқын статистикалық артықшылық байқалмады. Бұл pilot-экспериментте негізгі айырмашылықтардың дәлдік метрикаларында көрінгенін, ал уақытша тұрақтылық көрсеткіштері үшін неғұрлым күрделі мультимодальды тексеру қажет екенін көрсетеді.

Мәтіндік robustness test нәтижелері мәтінге бұзылу енгізілген сайын барлық модельдердің сапасы төмендейтінін көрсетті. Linear SVM үшін орташа UAR 0,254-тен 0,237 және 0,227-ге дейін төмендеді; logistic regression үшін 0,254-тен 0,232 және 0,232-ге дейін; MLP үшін 0,215-тен 0,202 және 0,203-ке дейін өзгерді. Бұл мәтіндік сигнал сапасының төмендеуі эмоцияларды тану сапасына тікелей әсер ететінін және сызықтық модельдердің шағын теңгерілген ішкі жиынтықта неғұрлым орнықты baseline ретінде қарастырылуы мүмкін екенін көрсетеді.

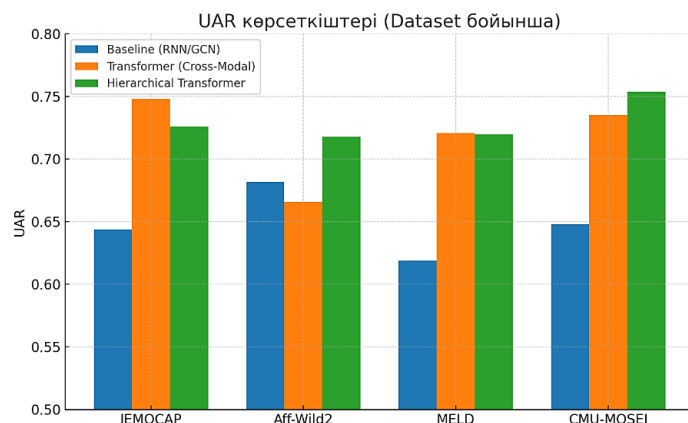
Практикалық тұрғыдан алғанда, эмоция болжамдарының уақытша тұрақтылығын бағалау адам-компьютер өзара әрекеттесу жүйелерінде маңызды рөл атқарады. Біріншіден, диалогтық интерфейстер мен аффективті агенттерде қысқа мерзімді жалған эмоциялық ауысуларды азайту жүйенің реакциясын неғұрлым сенімді және болжамды етуге мүмкіндік береді. Екіншіден, инклюзивті және көмекші жүйелерде эмоцияны уақыт бойынша дәйекті бағалау пайдаланушы күйін шамадан тыс ауытқуларсыз бақылауға көмектеседі. Осы тұрғыдан алғанда, ұсынылған тұрақтылық көрсеткіштері стандартты дәлдік метрикаларын толықтыратын қолданбалы бағалау құралы ретінде қарастырылуы мүмкін.

ҚОРЫТЫНДЫ

Бұл жұмыста эмоцияларды тану модельдерінің уақыт бойынша тұрақтылығын бағалауға арналған тәсіл қарастырылды. Негізгі назар стандартты дәлдік көрсеткіштерін уақытша дәйектілікті сипаттайтын қосымша метрикалармен толықтыруға аударылды. Осы мақсатта N_{switch} , D_{avg} және $AC(1)$ көрсеткіштері эмоция болжамдарының тұрақтылығын сипаттайтын негізгі құралдар ретінде ұсынылды, ал Херст көрсеткіші H ұзақ мерзімді тәуелділікті сипаттайтын қосымша теориялық метрика ретінде қарастырылды.

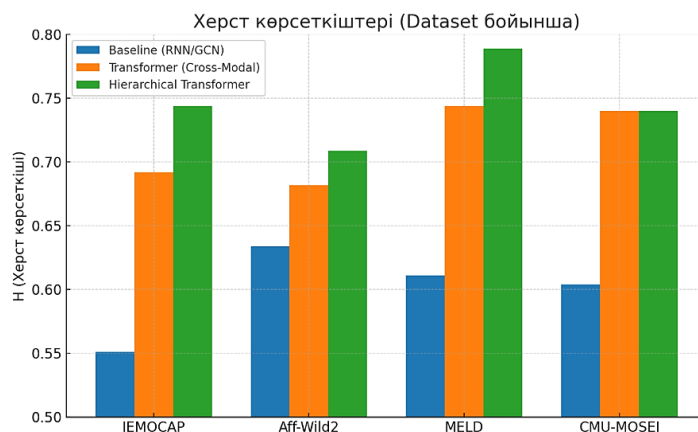
Әдебиеттерге және салыстырмалы талдауға сүйене отырып, трансформаторлық және иерархиялық архитектуралардың уақытша тәуелділіктерді модельдеуде перспективалы екені көрсетілді. Сонымен қатар, қосымша pilot-эксперимент ретінде MELD деректер жиынтығының теңгерілген шағын ішкі жиынтығында мәтіндік baseline-модельдер бағаланды. Бұл тәжірибеде linear SVM орташа сапа бойынша ең жақсы нәтижелерді көрсетті, logistic regression жақын baseline ретінде байқалды, ал MLP моделі дәлдік көрсеткіштері бойынша статистикалық мәнді түрде төмен нәтиже берді. Уақытша тұрақтылық көрсеткіштері бойынша модельдер арасында айқын статистикалық айырмашылық анықталған жоқ, бұл олардың күрделірек мультимодальды жағдайда әрі қарай тексерілуі қажет екенін көрсетеді.

Осылайша, жұмыстың негізгі нәтижесі — эмоцияларды тану жүйелерін бағалауда уақытша тұрақтылықты дербес және маңызды сипаттама ретінде қарастыру қажеттігін көрсету. Ұсынылған тәсіл эмоция болжамдарының тек дәлдігін ғана емес, уақыт бойынша дәйектілігін де бағалауға мүмкіндік береді. Болашақ зерттеулерде бұл талдау толық мультимодальды параметрлерде, соның ішінде аудио, бейне және мәтін арналарын бірлесіп пайдаланатын нақты модельдерде кеңейтілуі тиіс.



2-сурет. UAR көрсеткіштері (төрт деректер жиынтығы бойынша салыстыру)

Ескерту – авторлар алынған нәтижелер негізінде құрастырған.



3-сурет. Херст көрсеткіші (H) (уақытша тұрақтылықты бағалау)

Ескерту – авторлар алынған нәтижелер негізінде құрастырған

МҮДДЕЛЕР ҚАЙШЫЛЫҒЫ: Автор(лар) мүдделер қайшылығы жоқ екенін мәлімдейді.

ҚАРЖЫЛАНДЫРУ: Бұл жұмыс Қазақстан Республикасы Ғылым және жоғары білім министрлігі Ғылым комитетінің қаржылай қолдауымен орындалды (Грант № BR24992875).

ИНСТИТУЦИОНАЛДЫҚ ЭТИКА КОМИТЕТІНІҢ ҚОРЫТЫНДЫСЫ: Қолданылмайды.

ЕРІКТІ КЕЛІСІМ ТУРАЛЫ МӘЛІМДЕМЕ: ИНСТИТУЦИОНАЛДЫҚ ЭТИКА КОМИТЕТІНІҢ ҚОРЫТЫНДЫСЫ: Қолданылмайды.

ДЕРЕКТЕРДІҢ ҚОЛЖЕТІМДІЛІГІ ТУРАЛЫ МӘЛІМДЕМЕ: Осы зерттеудің нәтижелерін растайтын деректерді тиісті сұраныс бойынша жауапты автордан алуға болады.

АЛҒЫС БІЛДІРУ: Авторлар зерттеу жұмысына әдістемелік қолдау көрсеткен әріптестеріне және мақаланың сапасын арттыруға ықпал еткен құнды пікірлері үшін рецензенттерге алғыс білдіреді.

ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫН ПАЙДАЛАНУ ТУРАЛЫ ХАБАРЛАМА: Қолжазбаны дайындау барысында авторлар жасанды интеллект құралдарын мәтінді редакциялау, ғылыми стильді жақсарту және тілдік тұрғыдан тексеру мақсатында пайдаланды. Жасанды интеллект құралдары ғылыми нәтижелерді алу, деректерді талдау немесе қорытындыларды қалыптастыру үшін қолданылған жоқ.

ӘДЕБИЕТТЕР ТІЗІМІ

- Jayasinghe, H. M., Wong, K. W., & Nugaliyadde, A. (2026). A systematic review of interpretability and explainability for speech emotion features in automatic speech emotion recognition. *Pattern Recognition*, 171(Part A). <https://doi.org/10.1016/j.patcog.2025.112122>
- Wang, X., Cheng, Y., Cheng, P., & Zhang, X. (2026). AttDNet: An attribute-driven deep learning network for visual emotion analysis. *Pattern Recognition*, 171(Part A). <https://doi.org/10.1016/j.patcog.2025.112098>
- Zhang, X., Li, G., Li, X., Liang, B., & Chen, Y. (2026). Sarcasm detection enhanced by multi-modal topics using denoising diffusion probabilistic models. *Pattern Recognition*, 171(Part A). <https://doi.org/10.1016/j.patcog.2025.112130>
- Pradhan, A., & Srivastava, S. (2026). A novel hybrid primary capsule based convolutional neural network model for emotion classification from physiological signals. *Biomedical Signal Processing and Control*, 112, Article 108541. <https://doi.org/10.1016/j.bspc.2025.108541>
- Qi, Y., Huang, H., & Zhang, H. (2022). Research on speech emotion recognition method based A-CapsNet. *Applied Sciences*, 12(24), 12983. <https://doi.org/10.3390/app122412983>
- Van, L.-T., Nguyen, Q.-H., & Le, T.-D.-T. (2022). Emotion recognition with capsule neural network. *Computer Systems Science and Engineering*, 41(3), 1083–1098. <https://doi.org/10.32604/csse.2022.021635>
- Liu, Z.-T., Han, M.-T., Wu, B.-H., & Rehman, A. (2023). Speech emotion recognition based on convolutional neural network with attention-based bidirectional long short-term memory network and multi-task learning. *Applied Acoustics*, 202. <https://doi.org/10.1016/j.apacoust.2022.109178>
- Sun, L., Zou, B., Fu, S., Chen, J., & Wang, F. (2019). Speech emotion recognition based on DNN-decision tree SVM model. *Speech Communication*, 115, 29–37. <https://doi.org/10.1016/j.specom.2019.10.004>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). <https://doi.org/10.1109/CVPR.2016.90>
- Kiranyaz, S., Ince, T., Abdeljaber, O., Avci, O., & Gabbouj, M. (2019). 1-D convolutional neural networks for signal processing applications. In *ICASSP 2019 – IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 8360–8364). <https://doi.org/10.1109/ICASSP.2019.8682194>
- Ghosal, D., Majumder, N., Poria, S., Chhaya, N., & Gelbukh, A. (2019). DialogueGCN: A graph convolutional neural network for emotion recognition in conversation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. <https://doi.org/10.18653/v1/D19-1015>
- Majumder, N., Hazarika, D., Gelbukh, A., Cambria, E., & Poria, S. (2019). DialogueRNN: An attentive RNN for emotion detection in conversations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(1), 6818–6825. <https://doi.org/10.1609/aaai.v33i01.33016818>
- Tsai, Y.-H. H., Bai, S., Liang, P. P., Kolter, J. Z., Morency, L.-P., & Salakhutdinov, R. (2019). Multimodal transformer for unaligned multimodal language sequences. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 6558–6569). <https://doi.org/10.18653/v1/P19-1656>
- Li, J., Wang, S., Chao, Y., Liu, X., & Meng, H. (2022). Context-aware multimodal fusion for emotion recognition. In *Proceedings of Interspeech 2022* (pp. 2013–2017). <https://doi.org/10.21437/Interspeech.2022-10592>
- Moorthy, S., & Moon, Y.-K. (2025). Hybrid multi-attention network for audio–visual emotion recognition through multimodal feature fusion. *Mathematics*, 13(7), 1100. <https://doi.org/10.3390/math13071100>

- Hakim, A., Marsland, S., & Guesgen, H. W. (2013). Computational analysis of emotion dynamics. In 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII) (pp. 185–190). <https://doi.org/10.1109/ACII.2013.37>
- Plutchik, R. (1980). *Emotion: A psychoevolutionary synthesis*. Harper & Row.
- Busso, C., Bulut, M., Lee, C.-C., et al. (2008). IEMOCAP: Interactive emotional dyadic motion capture database. *Language Resources and Evaluation*, 42(4), 335–359. <https://doi.org/10.1007/s10579-008-9076-6>
- Kollias, D., Tzirakis, P., Nicolaou, M., et al. (2019). Aff-Wild2: Extended in-the-wild database for affect recognition. *ArXiv*. <https://doi.org/10.48550/arXiv.1811.07770>
- Poria, S., Hazarika, D., Majumder, N., et al. (2019). MELD: A multimodal multi-party dataset for emotion recognition in conversation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL)*. <https://doi.org/10.48550/arXiv.1810.02508>
- Liang, P. P., Zadeh, A., & Morency, L.-P. (2018). CMU-MOSEI: Multimodal opinion sentiment and emotion intensity. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (pp. 2236–2246). <https://doi.org/10.18653/v1/P18-1208>
- Zadeh, A., Chen, M., Poria, S., Cambria, E., & Morency, L.-P. (2017). Tensor fusion network for multimodal sentiment analysis. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1103–1114). <https://doi.org/10.18653/v1/D17-1115>
- Liu, Z., Shen, Y., Lakshminarasimhan, V., Liang, P. P., Zadeh, A., & Morency, L.-P. (2018). Low-rank multimodal fusion (LMF). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (pp. 2247–2256). <https://doi.org/10.18653/v1/P18-1209>
- Zadeh, A., Liang, P. P., Mazumder, N., Poria, S., Cambria, E., & Morency, L.-P. (2018). Memory fusion network for multi-view sequential learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1). <https://doi.org/10.1609/aaai.v32i1.12021>

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Кудубаева Сәуле Әлжанқызы – техника ғылымдарының кандидаты, қауымд. проф., Л.Н. Гумилев атындағы Еуразиялық ұлттық университеті, Астана қ., Қазақстан; жетекші ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы қ., Қазақстан

Кудубаева Сауле Альжановна – кандидат технических наук, ассоц.проф., Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Казахстан; ведущий научный сотрудник, «Институт информационных и вычислительных технологий» КН МНВО РК, г. Алматы, Казахстан

Saule Kudubayeva– Candidate of Technical Sciences, Associate Professor, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan; leading researcher, Institute of Information and computing technologies, Almaty, Kazakhstan
e-mail: saule.kudubayeva@gmail.com,
ORCID: <https://orcid.org/0000-0002-3807-9530>



Ернар Сексенбаев – докторант, Л.Н. Гумилев атындағы Еуразиялық ұлттық университеті, Астана қ., Қазақстан

Ернар Сексенбаев – докторант, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Казахстан

Yernar Seksenbayev – Doctoral student, Gumilyov Eurasian National University, Astana, Kazakhstan

e-mail: simpleuser6501@gmail.com

ORCID: <https://orcid.org/0000-0002-2863-1715>



Дайырбаева Эльмира Нурбеккызы – ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы қ., Қазақстан; магистр, аға оқытушы, Қ.И. Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы қ., Қазақстан

Дайырбаева Эльмира Нурбеккызы – научный сотрудник, «Институт информационных и вычислительных технологий» КН МНВО РК, г. Алматы, Казахстан; магистр, старший преподаватель, Казахский национальный исследовательский технический университет им. К.И. Сатпаева, г. Алматы, Казахстан

Elmira Daiyrbayeva – researcher, Institute of Information and computing technologies of the Ministry of education and science of the Republic of Kazakhstan, Almaty, Kazakhstan; master, senior lecturer, Satpayev University, Almaty, Kazakhstan

e-mail: nurbekkyzyelmira@gmail.com

ORCID: <https://orcid.org/0000-0002-4255-5456>



Еримбетова Айгерім Сембековна – PhD, қауымд. профессор, жетекші ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы қ., Қазақстан

Еримбетова Айгерим Сембековна – PhD, ассоц. профессор, ведущий научный сотрудник, «Институт информационных и вычислительных технологий» КН МНВО РК, г. Алматы, Казахстан

Aigerim Yerimbetova – PhD, Association professor, leading researcher, Institute of Information and computing technologies, Almaty, Kazakhstan

e-mail: aigerian8888@gmail.com

ORCID: <https://orcid.org/0000-0002-2013-1513>



Бурибаев Жолдас Алладинович – аға ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы, Қазақстан; PhD, Әл-Фараби атындағы қазақ Ұлттық Университеті, Алматы қ., Қазақстан

Бурибаев Жолдас Алладинович – старший научный сотрудник, «Институт информационных и вычислительных технологий» КН МНВО РК, г. Алматы, Казахстан; PhD, Казахский национальный университет им. Аль-Фараби, г. Алматы, Казахстан.

Zholdas Buribayev – Senior Researcher, Institute of Information and Computing Technologies Almaty, Kazakhstan; PhD, Al-Farabi Kazakh National University, Almaty, Kazakhstan.

e-mail: zholdas_87@mail.ru

ORCID: <https://orcid.org/0000-0002-3486-227X>