

Е.Т. Өмір¹, Ж.Д. Изтаев^{1*}, К.Т. Айхынбай¹, Ж.Ш. Аширбекова¹, З.К. Кушаров²

¹магистрант, М.Әуезов атындағы ОҚУ, Шымкент, Қазақстан

¹п.ғ.к., қауымдастырылған профессор, М.Әуезов атындағы ОҚУ, Шымкент, Қазақстан

¹магистр, оқытушы, М.Әуезов атындағы ОҚУ, Шымкент, Қазақстан

¹магистр, аға оқытушы, М.Әуезов атындағы ОҚУ, Шымкент, Қазақстан

²э.ғ.к., доцент, Ташкент мемлекеттік экономикалық университеті, Ташкент, Өзбекстан

*Корреспондент авторы: zhalgasbek71@mail.ru

TTS ТЕХНОЛОГИЯЛАРЫН ТАЛДАУ ЖӘНЕ LSTM НЕГІЗІНДЕГІ АДАПТИВТІ ИНТЕЛЛЕКТУАЛДЫ ОҚЫТУ ЖҮЙЕСІН ЖҮЗЕГЕ АСЫРУ

Түйін

Бұл мақалада мәтінді сөйлеуге түрлендіретін (Text-to-Speech, TTS) технологиялардың қазіргі жағдайы мен функционалдық мүмкіндіктері жан-жақты қарастырылды. Әртүрлі TTS сервистерінің дыбыстау сапасы, тілдік қолдауы, параметрлік икемділігі және техникалық интеграциялау мүмкіндіктері талданып, олардың білім беру саласында қолданылу әлеуеті көрсетілді. Жүргізілген талдау TTS технологияларының оқу материалдарын дыбыстау, ерекше білім беру қажеттіліктері бар пайдаланушыларды қолдау және цифрлық оқыту ортасын жетілдіру тұрғысынан маңызды рөл атқара алатынын дәлелдейді. Сонымен қатар жұмыста LSTM (Long Short-Term Memory) нейрондық желісінің архитектурасы мен жұмыс принциптері зерттелді. LSTM моделінің уақыттық тәуелділіктерді тиімді өңдеу қабілеті оның адаптивті оқыту жүйелерінде қолдануға қолайлы екенін негіздейді. Бұл модель оқушының оқу динамикасын талдауға, оның алдыңғы әрекеттерін ескеруге және жеке оқу траекториясын қалыптастыруға мүмкіндік береді. Жүргізілген зерттеу нәтижелері TTS пен LSTM технологияларын біріктіру заманауи, бейімделген және пайдаланушыға бағдарланған оқыту жүйелерін құрудың перспективалы бағыты екенін көрсетеді және бұл саладағы әрі қарайғы зерттеулерге негіз бола алады.

Кілттік сөздер: TTS, сөйлеу синтезі, LSTM, адаптивті оқыту, интеллектуалды оқыту жүйесі, бейімделген білім беру, нейрондық желі, мәтінді дыбысқа түрлендіру.

Кіріспе

Қазіргі білім беру жүйесінде ақпаратты өңдеу мен жеткізу тәсілдерінің әртүрлілігі маңызды рөл атқарады. Соңғы жылдары цифрлық технологиялардың дамуы білім беру процесін автоматтандыруға және жекелендірілген оқыту әдістерін енгізуге мүмкіндік берді. Солардың ішінде **сөйлеу синтезі (Text-to-Speech, TTS)** технологиялары ерекше орын алады.

TTS жүйелері мәтінді табиғи адам дауысына жақын дыбысқа түрлендіріп, оқу материалдарын тыңдау форматына ауыстыруға мүмкіндік береді. Бұл әсіресе оқушылардың оқу мотивациясын арттыруға, визуалды ақпаратқа шектеулі қол жеткізу жағдайында білім алуына және онлайн оқыту платформаларын жетілдіруге ықпал етеді.

Бұдан бөлек, адаптивті оқыту жүйелері оқу процесін жекелендіруді қамтамасыз етуі керек. Мұндай жүйелер пайдаланушының оқу қарқыны, алдыңғы білім деңгейі және оқу әрекеттерін ескере отырып, келесі оқу тапсырмаларын динамикалық түрде ұсынады. Осы мақсатқа жетуде **LSTM (Long Short-Term Memory)** нейрондық желілері маңызды құрал болып табылады [1].

Қазіргі таңда TTS технологиялары мен LSTM архитектурасын біріктіру арқылы **жекелендірілген және бейімделген интеллектуалды оқыту жүйесін** құру бағытында көптеген зерттеулер жүргізілуде. Мұндай жүйелер тек оқу материалын дыбыстаумен шектелмей, оқушының оқу тәжірибесін талдап, оқу траекториясын автоматты түрде қалыптастыра алады. Осы зерттеудің мақсаты – TTS сервистерінің функционалдық

мүмкіндіктерін талдау және LSTM моделін қолдану арқылы адаптивті интеллектуалды оқыту жүйесін жобалау.

Материалдар мен әдістер

Сөйлеу синтезі және мәтін анализі негізіндегі TTS жүйесі

Бұл кезеңде сервис мәтінді дыбысқа айналдыруға дайындайды. Мәтінді реттеу – сандар, аббревиатуралар мен таңбаларды анықтап, оларды мәтіндік формаға әкеледі. Мысалы, “56 см” – “елу алты сантиметр”. Лингвистикалық ретке келтіру – мәтінді мағыналық мазмұнға, сөз таптарына және екпіндерге бөледі. Фонемаларға бөлу – мәтіндегі сөздерді фонемаға, яғни дыбысқа айналдыру. Мысалы, “бала” сөзінде 4 әріп сәйкесінше 4 дыбыс – “/b/ /a/ /l/ /a/” бар. Просодиялық талдау – дыбысталу ерекшеліктерін, екпін, интонация, кідірістерді анықтайды. Бұл кезеңде сервис мәтінге сәйкес аудионы әзірлейді.

Акустикалық модель – Дыбысқа қажет параметрлерді жасайды: тон биіктігі, дыбыс ұзақтығы, амплитуда және т.б. Бұрындары қолданыста НММ болса, қазіргі таңда нейрондық желілер қолданылады: Tacotron, FastSpeech, VITS және т.б. Вокодер (Vocoder) – дыбыстау процессінің соңғы кезеңі. Дыбыстау параметрлерін нақты аудиосигналдарға айналдырады. Қазіргі заманғы вокодерлер ретінде WaveNet, WaveGlow, HiFi-GAN, MelGAN қолданыста бар.

TTS сервисінің жұмыс жасау процесі кезеңдерін математикалық (2) формула арқылы көрсетсек болады.

$$\hat{X} = \arg \max_x P(X | Y, \theta) \quad (2)$$

Мұндағы,

\hat{X} - синтезделген аудио;

$\arg \max_x$ — ықтималдықты максимумға жеткізетін X мәні; Y — кіріс мәтіні;

θ — модель параметрлері;

$P(X|Y, \theta)$ — мәтін Y және параметрлер θ берілгенде X аудиосигналын генерациялау ықтималдығы [2].

$\hat{X} = \arg \max_x P(X|Y, \theta)$ — мәтін Y және параметрлер θ берілгенде аудио сигнал X үшін шартты ықтималдықты ($P(X|Y, \theta)$) максимумға жеткізетін X мәні. Бұл формула TTS жүйесінің мақсатын көрсетеді: модель берілген мәтінге

(Y) және өзінің параметрлеріне (θ) сүйене отырып, соған сәйкес келетін ең шынайы және табиғи естілетін сөйлеу сигналын (X) генерациялайды. Яғни, барлық мүмкін сөйлеу нұсқаларының ішінен мәтінге ең жақсы сәйкес келетінін таңдайды. Мәтін алдымен фонетикалық немесе лингвистикалық өңдеуден өтеді, содан кейін модель оны мел-спектрограмма түріндегі дыбыстық бейнеге айналдырады. Бұл бейне вокодер деп аталатын бөлік арқылы нақты аудио сигналға түрлендіріледі. Нәтижесінде, TTS сервері пайдаланушы енгізген мәтінге сәйкес келетін сөйлеу файлы жасайды [3].

TTS жүйелерінің архитектурасы мен жұмыс принциптері қарастырылғаннан кейін, осы технологияны нақты жүзеге асыратын дайын сервистердің мүмкіндіктерін салыстыру маңызды. Бұл сервистер дыбыстау сапасы, қолдайтын тілдер жиынтығы, параметрлік икемділігі және техникалық интеграциялау тұрғысынан бір-бірінен ерекшеленеді. Төменде олардың негізгі сипаттамаларының жүйелі түрде салыстырылуы 6-кестеде ұсынылды.

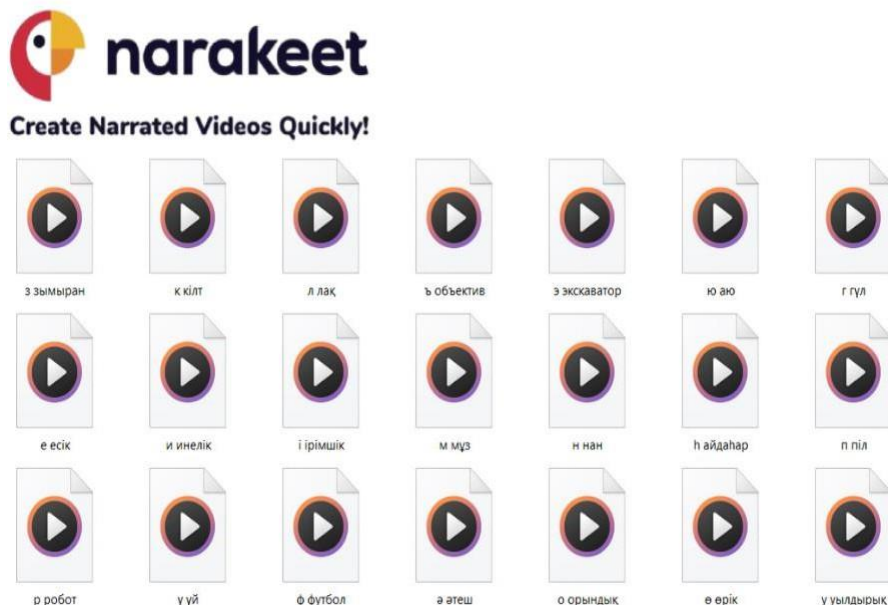
1 кесте. TTS сервисінің техникалық салыстырмалы кестесі

TTS сервер	Қазақ тілі	Дауыс модель саны	Дыбыс сапасы(MOS)	Sample Rate(дыбыстық жиілік)	Bitrate(Биттік жылдамдық)
Narakeet	Бар	900+	4.2-4.5	22050-44100 Hz	96–192 kbps (MP3)
Google TTS	Жоқ	220+	4.5-4.6	22050 Hz	Авто (әдетте 128 kbps)
Microsoft Azure	Жоқ	100+	4.2	24000 Hz	Авто
Amazon Polly	Шектеулі	60+	4.4-4.5	16000-22050 Hz	96–192 kbps

Осы салыстыру арқылы ең қолайлы және тиімді TTS сервисі таңдалып, жобаның талаптарына сәйкес интеграциялау шешімін қабылдау жүргізілді. Талдаулар жүргізу барысында көрсетілген бірнеше text-to-speech серверлерінің құрылымдық параметрлері мен сандық сипаттамалары ескерілді және ең қолайлы нұсқа ретінде Narakeet дыбыстау сервері таңдалды. Narakeet басқа серверлерге қарағанда дауыс моделінің көптігімен, соның ішінде қазақ тілін қолдау мүмкіндіктерімен және дыбыс сапасының жоғарылығымен ерекшеленді. Бұл TTS серверінің негізгі ерекшелігі – автоматты түрде мәтіннен дыбыс жасау жоғары деңгейде орындалады.

Narakeet - кез-келген сөзді, текстті автоматты түрде жасанды интеллекттің көмегімен дыбыстауға мүмкіндік беретін TTS(Text-to-Speech) сервер. Бұл сервер берілген тестті, сөзді дауыстық дыбыстап, mp3, m4a файл ретінде жібереді. Серверде 80-нен көп тілдік қолдаулар мен түрлі дауыстар қамтылған [4].

Narakeet дыбыстау ортасында әзірленген аудиофайлдар көрінісі 1-суретте көрсетілген.



150+ аудиофайл - генерацияланған суреттерге сәйкес дыбыстар генерацияланды.

1-сурет – Narakeet жасанды интеллектісі көмегімен әзірленген аудиофайлдар көрінісі

Narakeet дыбыстау жүйесінің техникалық баптауларын жүзеге асыру процесі:

1. Дыбысталуды қажет ететін сөздер тізімін әзірлеу- дыбыстау үшін кемінде 150 + сөз көлемінде сөздер жинақталды.

2. Жүйеден дауысты таңдау – қазақ тілін қолдайтын дауыстар ішінен ең үздік және жағымды естілетін нұсқа таңдап алынды.

3. Аудио-форматты таңдау: серверге жүктеуге ыңғайлы m4a форматы қолданылды.

4. Жүйеге интеграциялау: дайын дыбыстық жазбалар оқу материалдарымен біріктіріліп, серверге салынды [5].

Келесі 2-кестеде Narakeet TTS серверіне орнатылған негізгі параметрлердің сипаттамасы берілген. Бұл параметрлер генерация сапасын, дауыс тембрін, интонацияны және аудиофайлдың ұзақтығын реттеуге мүмкіндік береді.

2-кесте. Narakeet TTS серверінің параметрлері

Техникалық баптаулар	Опцияларды таңдау	Түсіндірме
Дауыстық дыбыстау	Cyrillic	Әйел дауысы
Дауыс жылдамдығы	1.0x	Қабылдауға ыңғайлы жылдамдық
Стиль	Conversational	Табиғи сөйлеу стилі
Дыбыс форматы	m4a, 192 kbps	Қалыпты аудио формат

Нәтижелер және талқылау

Берілген кесте мобильді қосымшада қолданылған дауыстық дыбыстауға қатысты техникалық баптауларды сипаттайды. Дауыстық дыбыстау бөлігінде мәтін кириллица қарпімен дыбысталып, әйел адамның дауысы таңдалған. Бұл балаларға жағымды және түсінікті үнмен ақпаратты жеткізу үшін маңызды. Дауыс жылдамдығы 1.0x болып орнатылған, бұл адам табиғи сөйлейтін қалыпты қарқын саналады және қабылдауға ыңғайлы. Стиль ретінде conversational, яғни табиғи сөйлеу стилі таңдалған. Мұндай стиль ресми емес, керісінше, достық қарым-қатынастағы әңгіме секілді естіледі, бұл балаларға материалды жеңіл қабылдауға көмектеседі. Дыбыс форматы ретінде m4a таңдалып, 192 kbps сапасында ұсынылған, бұл дыбыстың анық әрі жоғары сапалы болуын қамтамасыз етеді. Осы баптаулардың барлығы мобильді қосымшадағы дыбыстық мазмұнның сапалы әрі тартымды болуына бағытталған.

Қосымшамен байланысы

- Әрбір модуль сабақтарында сәйкес m4a – файлмен байланысқан;
- Объектке, яғни дыбыстық батырмаға басу арқылы активациялау жүзеге асырылған.

Тапсырмалар генерациясын әзірлеу үшін Chat GPT жасанды интеллектісі қолданылды.

GPT (Generative Pre-trained Transformer) - OpenAI компаниясы жасаған, табиғи тілді өңдеуге және генерациялауға арналған генеративті трансформерлік модель. GPT — автогрессионды тілдік модельдер қатарына жатады және мәтіндегі келесі токенді (сөз не сөз бөлігі) алдыңғы контекстке қарап болжауға үйренеді.

Трансформер архитектурасы

GPT тек трансформер архитектурасының декодер бөлігін пайдаланады. Бұл архитектура алғаш рет Vaswani және авторлар ұсынған "Attention is All You Need" атты ғылыми мақалада сипатталған. GPT моделінің негізінде self-attention (өзіне назар аудару) механизмі жатыр, ол модельге бүкіл контексті бір уақытта талдауға мүмкіндік береді.

Модельдің әр қабаты мына компоненттерден тұрады:

- көпбасты self-attention механизмі (multi-head self-attention),
- қабаттық нормализация (layer normalization),
- толық байланысқан нейрондық желі (feed-forward network),
- қалдық байланыстар (residual connections). GPT екі негізгі кезең бойынша оқытылады:

1. Алдын ала оқыту (Pre-training)

Бұл кезеңде модель үлкен көлемдегі мәтіндер негізінде келесі токенді болжауға үйренеді. Бұл - автогрессионды оқыту деп аталады. GPT интернеттен жиналған ашық дереккөздердегі мәтіндер арқылы тілдің құрылымын, мағынасын және логикасын үйренеді.

2. Қайта оқыту немесе нақты салаға бейімдеу (Fine-tuning) Бұл кезеңде модель белгілі бір тақырыпқа немесе мақсатқа бейімделіп, қосымша деректер арқылы оқытылады. Мысалы, балаларға арналған тапсырмалар генерациялау, медицина, программалау салалары. Мәтін генерациялау принципі

Оқыту аяқталған соң, модельге бастапқы мәтін (prompt) беріледі. Сол негізде ол біртіндеп келесі токендерді болжайды және бүкіл мәтінді құрастырады. Әр жаңа токен алдыңғы токендерге қарай ықтималдық бойынша таңдалады.

Формула түрінде былай беріледі (3):

$P(w_1, w_2, \dots, w_n) = \prod^n$	$P(w_t w_1, w_2, \dots, w_{t-1})$
-------------------------------------	-------------------------------------

Мұндағы,

- w_t - бұл t-нөмірлі токен, яғни мәтіндегі t-ші сөз (мысалы, 1-сөз, 2-сөз, 3-сөз, т.б.)
- $P(w_t | w_1, \dots, w_{t-1})$ - бұл t-нөмірлі сөздің ықтималдығы, алдыңғы (1-ден бастап t-1-ге дейінгі) сөздерге байланысты.

• Басқаша айтқанда: "осыған дейін не айтылды — соған қарап, келесі сөз қандай болуы мүмкін?"

• $\prod_{t=1}^n$

$P(w_t)$ - бұл математикалық түрде "көбейтінді" дегенді білдіреді.

Яғни барлық ықтималдықтарды бір-біріне көбейтеміз.

Бұл жасанды интеллект дипломдық жұмыста әртүрлі күрделілік деңгейіндегі тапсырмаларды құру үшін қолданылды. Осы мақсатта бірнеше GPT моделінің нұсқалары зерттеліп, олардың ерекшеліктері мен тиімділігі жан-жақты салыстырылды. Бұл салыстыру тапсырмалардың сапасын арттыруға және жобаға ең қолайлы модельді таңдауға мүмкіндік берді. Сондықтан әр нұсқаның техникалық мүмкіндіктері зерттеліп 3-кестеде салыстырылды.

3-кесте. Chat GPT-дің ескі және жаңа нұсқаларын салыстыру

Қасиет	GPT-2	GPT-3	GPT-4
Параметрлер саны	1.5 млрд	175 млрд	500+ млрд (жоғары болжам)
Контексті ұстау	1024 токен	2048 токен және одан жоғары	8000+ токен, кей нұсқаларда 32000 токенге дейін
Адаптация	Қарапайым баптау	Үлкен икемділік	Күрделі міндеттерге оңтайлы
Архитектура негізі	Трансформер декодер	Трансформер декодер	Трансформер декодер, жетілдірілген

Нейрондық қабаттар саны	48 қабат	96 қабат	100+ қабат (жоғары деңгейдегі архитектуралық жетілдірулер)
Қолдану шеңбері	Жалпы тілдік модельдер	Көп салалы, жоғары сапалы генерация	Көп салалы, жоғары дәлдік, креативтілік, көптілділік

Балаларға арналған тапсырмаларды автоматты түрде генерациялау мақсатында бірнеше GPT моделінің нұсқалары талданып, олардың техникалық сипаттамалары мен өнімділігі салыстырылды. Зерттеу нәтижесінде, ең жоғары дәлдік пен адаптивтілік көрсеткен, сонымен қатар күрделі мәтіндерді тиімді өңдеуге қабілетті GPT-4 моделі тапсырмаларды генерациялау үшін таңдалды.

Жобада балаларға арналған тапсырмалар екі негізгі принцип бойынша құрылды. Біріншісі – тапсырмаларды тақырыптық блоктарға бөлу, екіншісі – олардың қиындық деңгейі бойынша жіктелуі.

1. Тапсырмаларды тақырыптарға/тарауларға бөлу:

- Әріптер категориясы;
- Сандар;
- Қоршаған орта;
- Түстер;
- Пішіндер;
- Мамандықтар және т.б.

Тапсырмаларды категорияларға бөлу арқылы баланың әр-түрлі тақырыптар бойынша білім алуына және сол білімдерді бекітуге арналған тапсырмалар құрылды.

2. Тапсырмаларды деңгейлерге бөлу:

- 1-деңгей (score) – тапсырмалардың ең жеңіл нұсқалары
- 2-деңгей (score) - қиындығы жағынан орташа деңгейлі тапсырмалар
 - 3-деңгей (score) – бұл деңгейде ең күрделі тапсырмалар қамтылды.

Тапсырмаларды әзірлеуге дайындалған промттар мысалы 4-кестеде көрсетілген.

4-кесте. Тапсырмалар әзірлеуге арналған промттар

Тараулар	1 – деңгей. Жеңіл тапсырмалар	2 – деңгей. Орташа тапсырмалар	3 – деңгей. Күрделі тапсырмалар
Сандар	Цифрды тап! Цифрдың атауын тап!	Суреттегі заттарды сана!	Қарапайым қосу, азайту амалдарын орында!
Түстер	Түстің атын тап!	Заттың түсін анықта!	Пайда болатын түсті тап!
Логикалық ойлау	Кімге не сәйкес?	Көзбен қабылдауды дамыту!	Жетіспейтін бөлікті тап!
Мамандықтар	Мамандықты тап!	Мамандыққа тиесілі затты тап!	Суреттегі мамандық иесіне сәйкес сөйлемді толықтыр!

Кестеде ұсынылған тараулар мен олардың үш деңгейлі құрылымы негізінде тапсырмалар әзірленді. Әр деңгей progressively қиындайтын мазмұнмен құрылып, баланың ойлау, қабылдау және есте сақтау қабілеттерін дамытуға бағытталды. Жалпы алғанда, осы құрылым негізінде 130+ тапсырма жасалды [6].

Жүргізілген талдау нәтижесінде қазіргі заманғы TTS технологияларының негізгі артықшылықтары мен шектеулері анықталды. **Дыбыстау сапасы:** Tacotron, FastSpeech және VITS сияқты нейрондық желілер қолданылатын TTS сервистері табиғи адам дауысына жақын дыбысты қамтамасыз етеді. Алайда, интонациялық және эмоциялық нюанстар әлі де толық нақтыланбаған. **Тілдік қолдау:** Заманауи TTS жүйелері бірнеше тіл мен диалекті қолдай алады, бірақ кейбір сирек кездесетін тілдерде сапасы төмен болуы мүмкін. **Параметрлік икемділік:** Pitch, жылдамдық, дыбыс күшін реттеу мүмкіндігі бар, бұл жүйелерді білім беру материалдарын бейімдеу үшін тиімді етеді. **Техникалық интеграция:** API және SDK арқылы TTS сервистерін онлайн оқыту платформаларына оңай біріктіруге болады [7].

Бұл нәтижелер көрсеткендей, TTS технологиялары оқу материалдарын дыбыстау арқылы оқушылардың оқу процесіне белсенді қатысуына мүмкіндік береді, әсіресе визуалды ақпаратқа қол жеткізу қиын болған жағдайда.

Қорытынды

Бұл мақалада зерттеу нәтижелері көрсеткендей, TTS (Text-to-Speech) технологиялары мен LSTM (Long Short-Term Memory) нейрондық желілерін біріктіру білім беру саласында адаптивті, бейімделген және пайдаланушыға бағытталған оқыту жүйелерін құруға тиімді мүмкіндік береді.

TTS технологияларының мүмкіндіктері: Заманауи TTS жүйелері мәтінді табиғи адам дауысына жақын дыбысқа түрлендіре алады, оқу материалдарын тыңдау форматында ұсыну арқылы оқушылардың оқу мотивациясын арттырады, сондай-ақ ерекше білім беру қажеттіліктері бар пайдаланушыларға қолдау көрсетеді. **LSTM моделінің артықшылықтары:** Уақыттық реттіліктерді өңдеуге және ұзақ мерзімді тәуелділіктерді сақтауға қабілетті LSTM нейрондық желілері оқушының оқу динамикасын талдауға, оның алдыңғы әрекеттерін ескеруге және жеке оқу траекториясын құруға мүмкіндік береді.

TTS пен LSTM негізіндегі адаптивті интеллектуалды оқыту жүйелері заманауи білім беру процесінде маңызды рөл атқара алады және білім сапасын арттыруға үлес қосумен қатар жасанды интеллект саласындағы заманауи перспективалы бағыт болып табылады.

Әдебиеттер тізімі

1. Усен Е.Б. Применение нейросетевых моделей для семантического анализа публикаций в социальных сетях // Международный научный журнал «ВЕСТНИК НАУКИ», 2025, Т.4. №2(83), С. 480-485.
2. Захарова О. И. Семантический анализ и синтез текстовых данных // Вестник ВГУ. Серия: Системный анализ и информационные технологии, 2024, №4, С. 182-208.
3. Калижанова А., Маликова Ф., Дүйсенбек Ф., Дүйсенбек Н. Мәліметтерді өңдеудің модульдік жүйесін жобалаудың модельдері мен әдістерін зерттеу және құру // ҚазККА хабаршысы, 2023, Т.127. №4, Б. 351-358.
4. Баегизова А.С., Мырзабекова Г.Е., Алимагамбетова А.З., Мухамедрахимова Г.И., Кассим М. Интеллектуалды кластерлеу әдістерін қолдана отырып, қысқа мәтіндерді талдау // Международный журнал информационных и коммуникационных технологий, 2025, Т.6. №2, Б. 23-36.
5. Кайбасова Д.Ж., Маханова Б.М., Сүлеймен А.Е. Қазақ тіліндегі мәтін стилін анықтауда табиғи тілді өңдеу (NLP) әдістері // Университет еңбектері, 2021, №2, Б. 172-176.

6. Есенгабылов И., Алдабергенова А., Оразбаева А., Адамбекова А., Жапсарбаев Г. Анализ практического применения нейронных сетей для распознавания рукописного текста // Вестник КазАТК, №136(1), С.155–161.
7. Науменко А. О. Технология RAG (Retrieval-Augmented Generation) как инновационный подход в LLM // Вестник науки, 2025, Т. 5. №8(89), С. 280-289.

References

1. Usen E.B. Primenenie nejrosetevykh modelej dlja semanticheskogo analiza publikacij v social'nyh setjah // Mezhdunarodnyj nauchnyj zhurnal «VESTNIK NAUKI», 2025, Т.4. №2(83), S. 480-485.
2. Zaharova O. I. Semanticheskij analiz i sintez tekstovykh dannyh // Vestnik VGU. Seriya: Sistemnyj analiz i informacionnye tekhnologii, 2024, №4, С. 182-208.
3. Kalizhanova A., Malikova F., Dujsenbek F., Dyjsenbek N. Mәlimetterdi өңдеудің modul'dik zhыjesin zhubalaudың model'deri men әdisterin zertтеu zhөne қуру // ҚазККА habarshysy, 2023, Т.127. №4, В. 351-358.
4. Baegizova A.S., Myrzabekova G.E., Alimagambetova A.Z., Muhamedrahimova G.I., Kassim M. Intellectually klasterleu әdisterin қoldana otyrup, қысқа мәтіндерді талдау // Mezhdunarodnyj zhurnal informacionnyh i kommunikacionnyh tehnologij, 2025, Т.6. №2, В. 23-36.
5. Kajbasova D.Zh., Mahanova B.M., Sylejmen A.E. Қазак тилиндегі мәтін стилін анықтауда табиғи тилді өңдеу (NLP) әдистері // Университет еңбектері, 2021, №2, В. 172-176.
6. Esengabylov I., Aldabergenova A., Orazbaeva A., Adambekova A., Zhapsarbaev G. Analiz prakticheskogo primenenija nejronnyh setej dlja raspoznavanija rukopisnogo teksta // Vestnik KazATK, №136(1), S.155–161.
8. Naumenko A. O. Tekhnologiya RAG (Retrieval-Augmented Generation) kak innovacionnyj podhod v LLM // Vestnik nauki, 2025, Т. 5. №8(89), S. 280-289.

Е.Т. Омир¹, Ж.Д. Изтаев^{1*}, К.Т. Айхынбай¹, Ж.Ш. Аширбекова¹, З.К. Кушаров²

¹магистрант, ermahan.omir@mail.ru, ЮКУ им. М. Ауэзова, Шымкент, Казахстан
¹к.п.н., ассоциированный профессор, zhalgasbek71@mail.ru, ЮКУ им. М. Ауэзова, Шымкент, Казахстан

¹магистр, преподаватель, Akulaisha78@mail.ru, ЮКУ им. М. Ауэзова, Шымкент, Казахстан
¹магистр, старший преподаватель, ashirbekova7878@mail.ru, ЮКУ им. М. Ауэзова, Шымкент, Казахстан

²к.э.н., доцент, kusharovz@gmail.com, Ташкенский государственный экономический университет, Ташкент, Узбекистан

АНАЛИЗ ТЕХНОЛОГИЙ TTS И РЕАЛИЗАЦИЯ АДАПТИВНОЙ ИНТЕЛЛЕКТУАЛЬНОЙ ОБУЧАЮЩЕЙ СИСТЕМЫ НА ОСНОВЕ LSTM

Аннотация

В данной статье всесторонне рассмотрено текущее состояние и функциональные возможности технологий преобразования текста в речь (Text-to-Speech, TTS). Проанализированы качество озвучивания, языковая поддержка, параметрическая гибкость и возможности технической интеграции различных TTS-сервисов, а также показан их потенциал применения в сфере образования. Проведённый анализ доказывает, что технологии TTS могут играть важную роль в озвучивании учебных материалов, поддержке пользователей с особыми образовательными потребностями и совершенствовании цифровой образовательной среды. Кроме того, в работе исследованы архитектура и принципы работы нейронной сети LSTM (Long Short-Term Memory). Способность модели LSTM эффективно обрабатывать временные зависимости обосновывает её пригодность для использования в адаптивных обучающих системах. Эта модель позволяет анализировать динамику обучения учащегося, учитывать его предыдущие действия и формировать индивидуальную траекторию

обучения. Результаты проведённого исследования показывают, что интеграция технологий TTS и LSTM является перспективным направлением для создания современных, адаптивных и ориентированных на пользователя обучающих систем, а также может служить основой для дальнейших исследований в этой области.

Ключевые слова: TTS, синтез речи, LSTM, адаптивное обучение, интеллектуальная обучающая система, адаптивное образование, нейронная сеть, преобразование текста в речь.

Е.Т. Omir¹, Zh.D. Iztayev^{1*}, К.Т. Aikhynbay¹, Zh.Sh. Ashirbekova¹, Z.K. Kusharov²

¹Master's student, ermahan.omir@mail.ru, M. Auezov SKU, Shymkent, Kazakhstan

¹Candidate of Pedagogical Sciences, Associate Professor, zhalgasbek71@mail.ru, M. Auezov SKU, Shymkent, Kazakhstan

¹Master of Science, Lecturer, Akulaisha78@mail.ru, M. Auezov SKU, Shymkent, Kazakhstan

¹Master of Science, Senior Lecturer, ashirbekova7878@mail.ru, M. Auezov SKU, Shymkent, Kazakhstan

²Candidate of Economical Sciences, Associate Professor, kusharovz@gmail.com, Tashkent State University of Economics, Tashkent, Uzbekistan

ANALYSIS OF TTS TECHNOLOGIES AND IMPLEMENTATION OF AN ADAPTIVE INTELLIGENT LEARNING SYSTEM BASED ON LSTM

Abstract

This article provides a comprehensive review of the current state and functional capabilities of Text-to-Speech (TTS) technologies. The analysis covers speech quality, language support, parametric flexibility, and technical integration possibilities of various TTS services, as well as their potential applications in the field of education. The conducted analysis demonstrates that TTS technologies can play a significant role in narrating educational materials, supporting users with special educational needs, and enhancing digital learning environments. Furthermore, the study investigates the architecture and operating principles of the Long Short-Term Memory (LSTM) neural network. The LSTM model's ability to efficiently process temporal dependencies substantiates its suitability for use in adaptive learning systems. This model enables the analysis of students' learning dynamics, considers their previous actions, and facilitates the creation of individualized learning trajectories. The results of this study indicate that the integration of TTS and LSTM technologies represents a promising direction for the development of modern, adaptive, and user-centered learning systems and may serve as a foundation for further research in this area.

Keywords: TTS, speech synthesis, LSTM, adaptive learning, intelligent learning system, personalized education, neural network, text-to-speech conversion.