ТЕХНИКАЛЫҚ ҒЫЛЫМДАР
ТЕХНИЧЕСКИЕ НАУКИ
TECHNICAL SCIENCE

CENTRAL ASIAN SCIENTIFIC JOURNAL
ISSN 3135-3061
2026. №4 (32), 2-том. 219–222 б.
DOI: https://doi.org/10.5281/zenodo.24132385
Журнал сайты: https://cajournal.kz

ӘОЖ 004.89:37.0

ҮЛКЕН ТІЛДІК МОДЕЛЬДЕР: ЖҰМЫС ІСТЕУ ПРИНЦИПІ ЖӘНЕ БІЛІМ БЕРУДЕГІ ҚОЛДАНЫЛУЫ

К.И. Дулат
М.Х.Дулати атындағы университет

Дулат Камила Ильясқызы — Цифрлық технологиялар және жасанды интеллект, кафедрасының 4 курс студенті, М.Х.Дулати атындағы университет, Жамбыл облысы,Тараз қаласы.

Ғылыми жетекші: Шрымбай Дана Абилахатовна — Жаратылыстану ғылымдарының магистрі, Цифрлық технологиялар және жасанды интеллект, Кафедрасының аға оқытушысы, М.Х.Дулати атындағы университет, Жамбыл облысы,Тараз қаласы.

© К.И. Дулат, 2026

Аннотация: Соңғы жылдары жасанды интеллект саласында үлкен тілдік модельдер (LLM) ерекше орын алып отыр. Мақалада тілдік модельдердің n-граммнан Transformer архитектурасына дейінгі даму жолы, зейін (attention) механизмінің жұмыс принципі, модельдерді оқыту кезеңдері қарастырылған. Сонымен қатар информатиканы оқытуда үлкен тілдік модельдерді қолданудың мүмкіндіктері мен тәуекелдері талданып, оларды тиімді әрі саналы пайдалану бойынша ұсыныстар берілген.

Түйін сөздер: жасанды интеллект, үлкен тілдік модель, Transformer, зейін механизмі, ChatGPT, білім беру

Соңғы жылдары жасанды интеллект саласындағы ең елеулі жетістіктердің бірі – үлкен тілдік модельдердің (ағылш. Large Language Models, LLM) пайда болуы. Мұндай модельдер адам тілін түсіну, мәтін құрастыру, аудару, қысқаша мазмұндау және бағдарламалық код жазу сияқты көптеген міндетті бір жүйе аясында орындай алады. Олардың кең таралуы тек ақпараттық технологиялар саласына ғана емес, білім беруге, ғылыми зерттеулерге және бизнеске де әсер етіп отыр [1]. Сондықтан тілдік модельдердің қалай жұмыс істейтінін түсіну қазіргі информатика мамандарына ғана емес, оқытушылар мен студенттерге де қажетті дағдыға айналды.

Тілдік модель – сөздер тізбегінің ықтималдығын бағалайтын математикалық модель. Қарапайым тілмен айтқанда, ол берілген мәтіннен кейін қандай сөздің (дәлірек айтқанда, токеннің) келуі ықтимал екенін болжайды. Бұл идея жаңа емес: n-грамм модельдері бірнеше онжылдық бойы сөйлеуді тану мен машиналық аударма жүйелерінде қолданылды. Алайда мұндай модельдер алдыңғы бірнеше сөзді ғана ескереді, сондықтан мәтіннің бір-бірінен алшақ орналасқан бөліктерінің арасындағы байланысты көрсете алмайды [2].

Нейрондық желілердің дамуымен рекуррентті нейрондық желілер (RNN) және олардың жетілдірілген түрі – LSTM пайда болды. Олар мәтінді сөзбе-сөз, тізбектей өңдейтіндіктен, ұзын мәтіндерде бастапқы ақпарат жоғала бастады, ал есептеулерді параллель жүргізу қиындады. Бұл шектеулерді 2017 жылы ұсынылған Transformer архитектурасы шешті: онда мәтіннің барлық токендері бір мезгілде өңделеді, ал олардың

өзара байланысы «зейін» (attention) механизмі арқылы анықталады [3]. Екі архитектураның негізгі айырмашылықтары 1-кестеде берілген.

Кесте 1 – RNN / LSTM және Transformer архитектураларын салыстыру

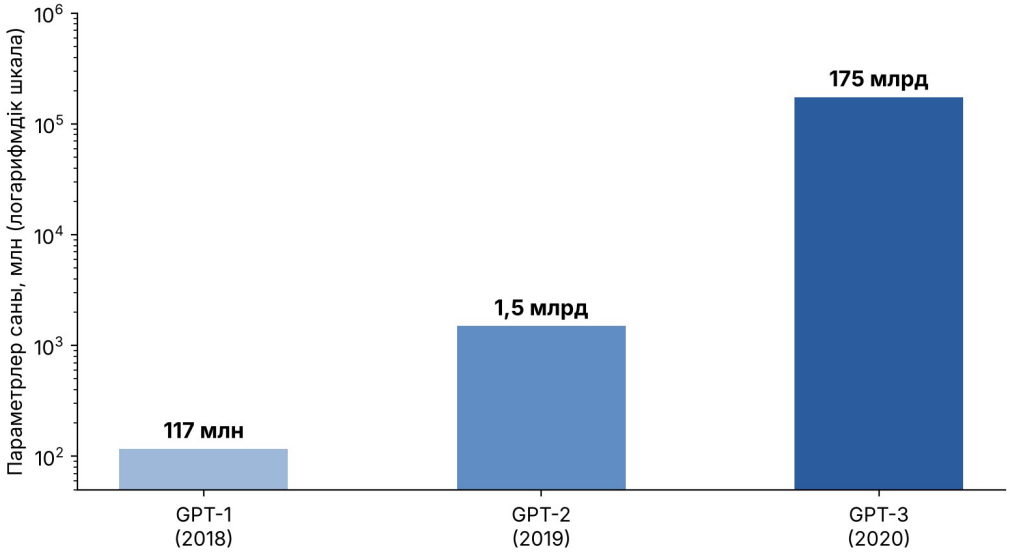
Көрсеткіш	**RNN / LSTM**	**Transformer**
Өңдеу тәсілі	Токендер тізбектей өңделеді	Барлық токендер бір мезгілде өңделеді
Алыс байланыстар	Ұзын мәтінде ақпарат жоғалады	Зейін арқылы тікелей есепке алынады
Параллель оқыту	Шектеулі	Жоғары
Масштабтау	Қиын	Салыстырмалы түрде оңай

Transformer-дің негізі – өзіндік зейін (self-attention) механизмі. Алдымен мәтін токендерге бөлінеді, әр токен сандық векторға – эмбеддингке айналдырылады, содан кейін оған сөздің мәтіндегі орнын көрсететін позициялық ақпарат қосылады. Әр токен үшін үш вектор – сұраныс (Q), кілт (K) және мән (V) есептеледі. Сұраныс пен кілттердің скалярлық көбейтіндісі қарастырылып отырған токеннің басқа токендермен қаншалықты байланысты екенін көрсетеді, ал нәтиже softmax функциясы арқылы салмақтарға айналып, мән векторларының өлшенген қосындысын құрайды [3]:

$$Attention(Q, K, V) = softmax(QK^T / \sqrt{d_k}) \cdot V \quad (1)$$

мұндағы d_k – кілт векторының өлшемі. Бұл формула модельге әрбір токенді өңдеу кезінде мәтіннің ең маңызды бөліктеріне «назар аударуға» мүмкіндік береді. Мысалы, «Студент бағдарламаны жазды, өйткені ол оны тапсыруы керек еді» сөйлемінде «ол» есімдігінің қай сөзге қатысты екенін модель зейін салмақтары арқылы анықтай алады.

Үлкен тілдік модельдерді даярлау әдетте екі кезеңнен тұрады. Бірінші кезеңде – алдын ала оқытуда – модель кітаптардағы, веб-беттердегі және басқа да көздердегі өте үлкен мәтін жиынтығында келесі токенді болжауға үйретіледі. Модельдердің мүмкіндіктері көбіне олардың көлемімен – параметрлер санымен – байланысты өсті. GPT отбасының алғашқы нұсқасында шамамен 117 млн параметр болса [4], GPT-2-де 1,5 млрд [5], ал 2020 жылы ұсынылған GPT-3-те 175 млрд параметр болды [6] (1-сурет).



Сурет 1 – GPT отбасы модельдеріндегі параметрлер санының өсуі (2018–2020 жж.)

Екінші кезеңде модель нақты міндеттерге бейімделеді. Атап айтқанда, ол адамдардың нұсқауларын орындауға және жауаптарды бағалау нәтижелері негізінде жақсартуға үйретіледі (RLHF әдісі), бұл модельді қарапайым мәтін жалғастырушыдан диалогқа қабілетті көмекшіге айналдырады [7]. Жұмыс кезінде модель жауапты токен-токенімен жасайды: әр қадамда келесі токеннің ықтималдықтар үлестірімі есептеліп, сол үлестірімнен таңдау

жүргізіледі. Сондықтан жауаптың сипаты генерация параметрлеріне де тәуелді. Мысалы, «температура» параметрі төмен болғанда модель ең ықтимал токендерді таңдап, жауап тұрақты әрі болжамды болады, ал жоғары болғанда жауап әртүрлі, шығармашылық сипатта болады, бірақ қателік ықтималдығы да артады.

Модель жауабының сапасы сұраныстың (промпттың) сапасына да байланысты. Нақты міндет қою, контекст беру, жауап форматын көрсету және мысал келтіру нәтижені жақсартады. GPT-3 туралы жұмыста бірнеше мысал беру арқылы модельдің жаңа міндетті қосымша оқытусыз орындай алатыны көрсетілген [6]. Бұл тәсіл кейін «промпт-инженерия» деп аталды.

Үлкен тілдік модельдердің білім берудегі мүмкіндіктері ерекше назар аударарлық. Біріншіден, олар студентке күрделі тақырыпты өз деңгейіне сай түсіндіріп, мысалдар келтіре алады, яғни жеке көмекші-репетитор рөлін атқарады. Екіншіден, информатика пәндерінде модель бағдарлама кодындағы қателерді табуға, алгоритмді түсіндіруге, жаттығулар мен тест тапсырмаларын құрастыруға көмектеседі. Үшіншіден, оқытушылар үшін оқу материалдарын дайындауға кететін уақыт қысқарады [8]. Информатика пәнінде қолдануға болатын кейбір мысалдар 2-кестеде берілген. Сонымен қатар көптілді модельдер қазақ тіліндегі оқу материалдарын жасауға да мүмкіндік береді, дегенмен қазақ тілі сияқты ресурстары шектеулі тілдер үшін жауап сапасы ағылшын тіліне қарағанда төмен болуы мүмкін.

Кесте 2 – Информатиканы оқытуда үлкен тілдік модельдерді қолдану мысалдары

Міндет	**Сұраныс мысалы**	**Нәтижені тексеру**
Алгоритмді түсіндіру	«Екілік іздеу алгоритмін 1 курс студентіне қарапайым тілмен түсіндір»	Оқулықпен салыстыру
Қатені табу	«Осы Python кодындағы қатені тауып, себебін түсіндір»	Кодты іске қосып тексеру
Тапсырма құрастыру	«Циклдер тақырыбы бойынша 5 есеп құрастыр»	Оқытушының қарап шығуы

Сонымен бірге бұл технологияның елеулі тәуекелдері бар. Ең алдымен, модель сенімді естілетін, бірақ шындыққа жанаспайтын жауап беруі мүмкін («галлюцинация»), өйткені ол фактіні тексермейді, тек ықтимал мәтінді жасайды. Оқыту деректеріндегі қателіктер мен біржақтылық модель жауаптарында да көрініс табуы мүмкін, ал өте үлкен модельдерді даярлау айтарлықтай есептеу ресурстары мен энергияны қажет етеді [9]. Білім беруде басты қауіп – студенттің тапсырманы өзі орындамай, дайын жауапқа сүйенуі және академиялық адалдық мәселесі [8].

Қорытындылай келе, үлкен тілдік модельдер – Transformer архитектурасы, өте үлкен деректер жиынтығы және есептеу қуатының үйлесімі нәтижесінде пайда болған құрал. Олар білім беру үдерісін айтарлықтай жеңілдете алады, бірақ тек саналы әрі сыни тұрғыда қолданылғанда ғана тиімді. Тәжірибеде мынадай қағидаларды ұстану ұсынылады: модель берген жауапты міндетті түрде басқа сенімді дереккөздермен тексеру; модельді ойлауды алмастыратын емес, оған көмектесетін құрал ретінде пайдалану; оқу орнында жасанды интеллектті қолдану ережелерін алдын ала анықтау. Тілдік модельдердің жұмыс принциптерін түсіну болашақта информатика мамандарының негізгі құзыреттерінің бірі болады деп күтуге болады.

Пайдаланылған әдебиеттер:

1. Russell S., Norvig P. Artificial Intelligence: A Modern Approach. 4th ed. – Hoboken: Pearson, 2020. – 1136 p.

2. Jurafsky D., Martin J.H. Speech and Language Processing (3rd ed. draft) [Электронный ресурс]. URL: https://web.stanford.edu/~jurafsky/slp3/ (дата обращения: 11.11.2026).

3. Vaswani A., Shazeer N., Parmar N. et al. Attention Is All You Need // Advances in Neural Information Processing Systems 30 (NIPS 2017). – 2017. – P. 5998–6008.

4. Radford A., Narasimhan K., Salimans T., Sutskever I. Improving Language Understanding by Generative Pre-Training. – San Francisco: OpenAI, 2018.

5. Radford A., Wu J., Child R. et al. Language Models are Unsupervised Multitask Learners. – San Francisco: OpenAI, 2019.

6. Brown T.B., Mann B., Ryder N. et al. Language Models are Few-Shot Learners // Advances in Neural Information Processing Systems 33 (NeurIPS 2020). – 2020. – P. 1877–1901.

7. Ouyang L., Wu J., Jiang X. et al. Training language models to follow instructions with human feedback // Advances in Neural Information Processing Systems 35 (NeurIPS 2022). – 2022. – P. 27730–27744.

8. Kasneci E., Sessler K., Küchemann S. et al. ChatGPT for good? On opportunities and challenges of large language models for education // Learning and Individual Differences. – 2023. – Vol. 103. – Art. 102274.

9. Bender E.M., Gebru T., McMillan-Major A., Shmitchell S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? // Proceedings of FAccT '21. – 2021. – P. 610–623.

© Central Asian Scientific Journal, 2026