15-дәріс Интеллектуалды ақпараттық жүйелердің дамуындағы заманауи үрдістер

Пәні: Интеллектуалды ақпараттық жүйелер

Алдыңғы тақырып: Көпагентті интеллектуалды жүйелер

Бұл дәріс интеллектуалды ақпараттық жүйелердің қазіргі архитектуралық, модельдік және басқарушылық бағыттарын түсіндіреді. Материал 25 слайдпен бірдей ретте құрылған және оқытушының бірінші жақтағы баяндауына бейімделген.

Дәрістің мақсаты

Студенттерге интеллектуалды ақпараттық жүйелердің заманауи даму бағыттарын, олардың формулалық негіздерін, практикалық пайдасын, шектеулерін және жауапты енгізу қағидаларын жүйелі түрде түсіндіру.

Күтілетін оқу нәтижелері

- cloud, edge AI, IoT және lakehouse архитектураларын салыстырады
- іргелі, генеративті және мультимодальды модельдердің рөлін түсіндіреді
- RAG, агенттік AI, AutoML және MLOps логикасын түсіндіреді
- XAI, fairness, federated learning және Green AI формулаларын талдайды
- қауіпсіздік, интероперабельдік және тәуекелге негізделген басқаруды негіздейді
- нақты ұйымға технологиялық даму жол картасын ұсынады

Заманауи үрдістер картасы

Бағыт	Негізгі мақсат	Пайда	Бақылау сұрағы
Edge AI	Дерекке жақын есептеу	Аз кідіріс	Құрылғы ресурсы жеткілікті ме?
RAG	Сенімді білімді қосу	Дереккөзді жауап	Құжат өзекті ме?
MLOps	Өмірлік циклді басқару	Қайталанғыштық	Drift бақылана ма?
Responsible AI	Зиянды азайту	Сенімді қолдану	Адам бақылауы бар ма?
Green AI	Ресурсты үнемдеу	Төмен шығын	Сапа өсімі энергияға тұра ма?

Заманауи үрдістерге кіріспе

Алдымен тақырыптың жалпы мағынасын ашамын.

Сәлеметсіздер ме! Бүгін мен интеллектуалды ақпараттық жүйелердің дамуындағы заманауи үрдістерді қарастырамын. Біз жеке технологияларды жаттап қана қоймай, олардың ұйымдағы дерек, шешім және адам еңбегін қалай өзгертетінін байланыстырамыз.

Қазіргі интеллектуалды жүйе бір модельден тұрмайды. Ол дерек платформасын, машиналық оқытуды, бұлтты немесе шеткі есептеуді, интерфейсті, қауіпсіздік пен басқаруды біріктіреді. Сондықтан үрдісті бағалағанда оның пайдасын, шектеуін және енгізу шартын бірге қарау керек.

Қарапайым мысал. Университеттегі кеңесші жүйе студент сұрағын түсінеді, оқу жоспарын дерекқордан табады, жауап жобасын құрады және күмәнді жағдайда маманға жібереді. Мұнда NLP, іздеу, генеративті модель және адам бақылауы бірге жұмыс істейді.

Негізгі ой: заманауи интеллектуалды жүйенің құндылығы бір алгоритмде емес, технологиялар мен жауапты басқарудың тұтас байланысында.

Дәрістің мақсаты және оқу нәтижелері

Енді оқу мақсатымызды нақтылап алайық.

Алдымен біздің оқу бағытымызды нақтылаймын. Мен бұлт пен edge AI-дан бастап іргелі модельдерге, мультимодальдыққа, RAG, агенттік жүйелер, MLOps, цифрлық егіздер және жауапты AI қағидаларына дейін бірізді өтемін.

Дәріс соңында сіздер негізгі үрдістерді ажыратып, әрқайсысының не үшін қажет екенін түсіндіре аласыздар. Сонымен қатар архитектура таңдау, модель сапасын өлшеу, әділдік, құпиялық, энергия тиімділігі және тәуекелді басқару арасындағы байланысты талдайсыздар.

Қарапайым мысал. Аурухана жаңа AI қызметін таңдағанда тек дәлдікке қарамайды. Ол жауап уақыты, деректің құпиялығы, түсіндірмелілік, қате бағасы және дәрігердің шешімдегі рөлін де бағалайды.

Негізгі ой: технологиялық үрдісті дұрыс түсіну үшін мүмкіндік пен жауапкершілікті бір өлшемде қарастырамыз.

Ережелік жүйеден дерекке негізделген жүйеге

Осыдан кейін интеллектуалды жүйелердің эволюциясына көшемін.

Интеллектуалды жүйелердің алғашқы буынында білім көбіне IF–THEN ережелерімен жазылды. Қазір ережелерге статистикалық модель, нейрондық желі және үлкен көлемдегі дерек қосылды. Бұл жүйені икемдірек етеді, бірақ нәтижені тексеру міндетін күшейтеді.

Ережелік тәсіл түсінікті және тұрақты, ал машиналық оқыту жасырын заңдылықты деректен табады. Қазіргі гибридті архитектурада міндетті шектеулер ережемен, күрделі тану міндеті модельмен орындалады. Осылайша жүйе бақылауды сақтап, деректің күшін қолданады.

Қарапайым мысал. Несие өтінімінде жас шектеуі мен құжаттың толықтығы ережемен тексеріледі, ал төлем қабілеттілігінің тәуекелі тарихи дерекпен үйретілген модель арқылы бағаланады. Соңғы шешімге адам сарапшы қатысады.

Негізгі ой: заманауи жүйе ережені жоймайды, оны дерекке негізделген модельмен дұрыс біріктіреді.

Бұлттық және бұлтқа тәуелсіз архитектура

Енді архитектуралық өзгерісті қарастырамын.

Бұлттық есептеу интеллектуалды қызметті жылдам іске қосуға, есептеу қуатын сұранысқа қарай өзгертуге және модельді орталықтан жаңартуға мүмкіндік береді. Сонымен бірге ұйым бір провайдерге тәуелділік, желі үзілуі және дерек орналасуы сияқты мәселелерді алдын ала бағалайды.

Қазіргі үрдіс cloud-native қағидаларын қолданады: қызметтер контейнерде орындалады, API арқылы байланысады және жүктемеге қарай масштабталады. Бірақ барлық міндетті бұлтқа көшіру дұрыс емес. Деректің сезімталдығы мен кідіріс талабы архитектураны анықтайды.

Қарапайым мысал. Интернет-дүкен ұсыныс моделін бұлтта іске қосады, себебі сұраныс мереке кезінде күрт өседі. Ал төлем терминалы желі болмаса да негізгі қауіпсіздік тексерісін жергілікті орындайды.

Негізгі ой: бұлт икемділік береді, ал дұрыс архитектура бұлттың пайдасын тұрақтылық және тәуелсіздік талаптарымен теңестіреді.

Edge AI және жауап уақыты

Осы түсініктен кейін практикалық механизмге көшемін.

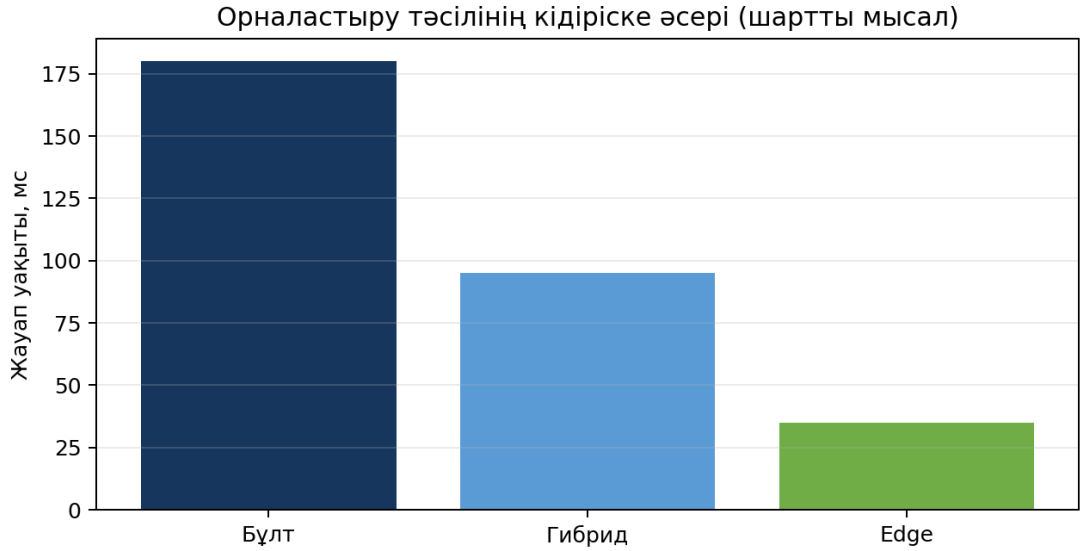
Edge AI модельді дерек пайда болатын құрылғыға немесе оған жақын серверге орналастырады. Бұл бейне, дыбыс және сенсор ағынын алыстағы бұлтқа толық жібермей, тез өңдеуге мүмкіндік береді.

Жалпы кідірісті $L_{total}=L_{sensor}+L_{network}+L_{compute}+L_{action}$ деп бөлуге болады. Edge тәсілі желілік бөлікті азайтады, бірақ құрылғы жадысы мен энергиясы шектеулі. Сондықтан модельді қысқарту, quantization және distillation тәсілдері маңызды болады.

Қарапайым мысал. Өндірістегі камера ақаулы бөлшекті 40 миллисекунд ішінде анықтауы керек. Егер сурет бұлтқа жіберілсе, желі кідірісі конвейерді тоқтатуға кедергі келтіруі мүмкін. Жергілікті модель шешімді бірден береді, ал статистика кейін бұлтқа жіберіледі.

Негізгі ой: edge AI жылдамдық пен құпиялықты арттырады, бірақ модель көлемі мен құрылғы ресурсын мұқият теңестіруді қажет етеді.

Ltotal = Lsensor + Lnetwork + Lcompute + Laction



1-сурет. Орналастыру тәсілі мен жауап уақытының шартты салыстырмасы

IoT және киберфизикалық жүйелер

Бұл үрдіс алдыңғы мүмкіндікті толықтырып, жаңа талап қояды.

Заттар интернеті сенсорлар мен құрылғыларды бір ақпараттық ортаға біріктіреді. Интеллектуалды ақпараттық жүйе осы ағыннан жағдайды бағалап, болжам жасап, физикалық нысанға әрекет командасын береді.

Киберфизикалық жүйеде қате тек экрандағы санға емес, нақты құрылғыға әсер етуі мүмкін. Сондықтан сенсор сапасы, уақыт белгісі, қауіпсіз байланыс және fail-safe режимі негізгі талапқа айналады.

Қарапайым мысал. Ақылды жылыжай температура, ылғал және жарықты өлшейді. Модель өсімдік күйін болжайды, ал басқару жүйесі суару уақытын реттейді. Сенсор қате мән берсе, жүйе бірден әрекет етпей, оны көрші сенсормен салыстырады.

Негізгі ой: IoT интеллектуалды жүйені нақты ортамен байланыстырады, сондықтан дерек сапасы мен физикалық қауіпсіздік бірге қаралады.

Үлкен деректер және lakehouse тәсілі

Енді мәселені жүйелік және жауапты қолдану тұрғысынан тереңдетемін.

Үлкен деректер тек көлемді емес, жылдамдық, әртүрлілік және сенімділік мәселесін қамтиды. Қазіргі жүйелер кестелік дерекпен қатар мәтін, сурет, журнал, сенсор және граф құрылымдарын өңдейді.

Data lake икемді шикі деректі сақтайды, data warehouse басқарылған талдауға ыңғайлы. Lakehouse осы екі тәсілдің артықшылығын біріктіруге тырысады: ашық файлдық сақтау, транзакция тәртібі, каталог және сапа бақылауы.

Қарапайым мысал. Сауда желісі чек, мобильді қосымша, қойма және пікір мәтіндерін бір ортада сақтайды. Сұраныс болжамы үшін барлық дерек бір клиент пен бір тауар анықтамасына дұрыс байланыстырылуы керек.

Негізгі ой: үлкен деректің құндылығы көлемде емес, сапалы каталог, ортақ мағына және қайталанатын өңдеу үдерісінде.

Іргелі модельдер және қайта қолдану

Енді алдыңғы ойды келесі технологиялық бағытпен байланыстырамын.

Іргелі модель немесе foundation model үлкен және әртүрлі деректе алдын ала үйретіліп, кейін көптеген міндетке бейімделеді. Бұл әр жобаға модельді нөлден үйрету қажеттігін азайтады.

Бейімдеу prompt, fine-tuning, adapter немесе сыртқы біліммен толықтыру арқылы орындалады. Үлкен модельдің мүмкіндігі жоғары, бірақ есептеу құны, жабық дерек, қате жауап және лицензия шарты есепке алынуы керек.

Қарапайым мысал. Қазақ тіліндегі өтініштерді жіктеу үшін ұйым жалпы тілдік модельді пайдаланып, шағын тексерілген мысалдармен бейімдейді. Бірақ құпия өтініштерді сыртқы сервиске жібермей, рұқсат етілген ортада өңдейді.

Негізгі ой: іргелі модель әмбебап негіз береді, ал нақты жүйе оны ұйымның дерегі, тілі және тәуекелімен бейімдейді.

Аудиторияға сұрақ: үлкен іргелі модельді әр міндетте қолдану міндетті ме?

Генеративті жасанды интеллект

Осы түсініктен кейін практикалық механизмге көшемін.

Генеративті AI жаңа мәтін, сурет, код, дыбыс немесе құрылымдық дерек жасайды. Оның басты айырмасы дайын санатты таңдаумен шектелмей, ықтимал мазмұн құрастыруында.

Тілдік модель келесі токеннің ықтималдығын P(x_t|x_1,…,x_{t-1}) бойынша бағалайды. Жауап табиғи көрінгенімен, модель фактіні міндетті түрде тексермейді. Сондықтан генерацияны дереккөз, шектеу және тексеру кезеңімен байланыстырамыз.

Қарапайым мысал. Оқытушы модельден зертханалық жұмыс жобасын алады, бірақ мақсат, термин, есеп және қауіпсіздік нұсқауын өзі тексереді. Модель көмекші болады, дайын жауаптың жалғыз авторы емес.

Негізгі ой: генеративті AI мазмұн жасауды жылдамдатады, ал факт пен жауапкершілік үшін тексерілетін жұмыс үдерісі қажет.

$$P(x_t \mid x_1, \ldots, x_{t-1})$$

Мультимодальды интеллектуалды жүйелер

Бұл үрдіс алдыңғы мүмкіндікті толықтырып, жаңа талап қояды.

Мультимодальды жүйе мәтін, сурет, дыбыс, бейне және сенсор мәліметін бір міндетте байланыстырады. Адам әлемді бірнеше арна арқылы қабылдайтындықтан, мұндай жүйе контексті толығырақ түсіне алады.

Модальдықтарды early fusion арқылы бастапқы белгілерде, late fusion арқылы бөлек модель нәтижелерінде немесе ортақ embedding кеңістігінде біріктіруге болады. Бір арна жоқ немесе сапасыз болса, жүйе сенімділікті төмендетуі тиіс.

Қарапайым мысал. Техник жабдықтың суретін көрсетіп, дауыспен «мына дыбыс неге өзгерді?» деп сұрайды. Жүйе бейнедегі бөлшекті, дыбыс спектрін және қызмет көрсету журналын бірге талдайды.

Негізгі ой: мультимодальдық көбірек контекст береді, бірақ әр дерек арнасының сапасы мен уақыт сәйкестігін бақылау қажет.

RAG және ұйым білімін пайдалану

Енді мәселені жүйелік және жауапты қолдану тұрғысынан тереңдетемін.

Retrieval-Augmented Generation немесе RAG сұраққа жауап берер алдында сенімді құжаттарды іздеп, табылған үзіндіні генеративті модельге контекст ретінде береді. Бұл модель білімін ұйымның өзекті дерегімен толықтырады.

Мәтіндер embedding векторына айналады. Ұқсастық $\cos(q,d)=(q\cdot d)/(\|q\|\cdot\|d\|)$ арқылы бағалануы мүмкін. Бірақ ең жақын үзінді әрқашан дұрыс емес, сондықтан дерекке қолжетімділік, жаңалық мерзімі, reranking және сілтеме көрсету қажет.

Қарапайым мысал. Студент «академиялық демалысқа қалай шығамын?» деп сұрайды. Жүйе ресми ереженің соңғы нұсқасын табады, тармақты көрсетеді және жауапты сол мәтінге сүйеніп құрастырады.

Негізгі ой: RAG модель жауабын тексерілетін ұйымдық дерекпен байланыстырады, бірақ іздеу сапасы жауап сапасының бөлігі болады.

$$\cos(q,d) = (q\cdot d)/(\|q\|\|d\|)$$

Аудиторияға сұрақ: іздеу дұрыс құжатты таппаса, генерация сапасына не болады?

Агенттік AI және құрал қолдану

Енді алдыңғы ойды келесі технологиялық бағытпен байланыстырамын.

Агенттік AI мақсатты бірнеше қадамға бөліп, жоспар құрып, құрал шақырып, нәтижені тексеріп, келесі әрекетті таңдайды. Бұл қарапайым чаттан күрделі жұмыс үдерісіне өтуді көрсетеді.

Агентке рұқсат шекарасы, әрекет журналы, уақыт және шығын лимиті қажет. Жоғары әсерлі операциялар адам бекітуінсіз орындалмауы тиіс. Автономдылық бақылаусыз әрекет дегенді білдірмейді.

Қарапайым мысал. Сатып алу агенті қойма қалдығын тексереді, үш жеткізушінің бағасын салыстырады және тапсырыс жобасын дайындайды. Бірақ төлемді жіберу алдында жауапты қызметкерге бекітуге ұсынады.

Негізгі ой: агенттік жүйе модельді әрекет ететін көмекшіге айналдырады, сондықтан рұқсат пен бақылау архитектураның негізгі бөлігі болады.

Аудиторияға сұрақ: агентке қандай әрекеттерді адам бекітуінсіз беруге болмайды?

AutoML және MLOps

Осы түсініктен кейін практикалық механизмге көшемін.

AutoML белгі таңдау, алгоритм салыстыру және гиперпараметр іздеу кезеңдерін автоматтандырады. MLOps модельді тәжірибеден өндірістік қызметке дейін басқаруға көмектеседі.

MLOps дерек нұсқасын, кодты, модельді, метриканы және орналастыру тарихын байланыстырады. Өндірісте accuracy жеткіліксіз: latency, drift, қате топтары және қызмет қолжетімділігі бақыланады.

Қарапайым мысал. Сұраныс болжау моделі ай сайын қайта үйретіледі. Жаңа нұсқа бірден барлық дүкенге берілмей, шағын топта A/B тексеріледі. Нәтиже нашарласа, жүйе алдыңғы нұсқаға автоматты қайтады.

Негізгі ой: AutoML модель құруды жылдамдатады, ал MLOps оның өмірлік циклін қайталанатын және бақыланатын етеді.

Цифрлық егіздер және симуляция

Бұл үрдіс алдыңғы мүмкіндікті толықтырып, жаңа талап қояды.

Цифрлық егіз нақты нысанның немесе үдерістің дерекпен жаңарып отыратын виртуалды үлгісі. Ол «не болып жатыр?» деген сұрақпен бірге «егер параметрді өзгертсек не болады?» деген сұраққа жауап береді.

Егіз сенсор ағынын, физикалық заңды, статистикалық модельді және сценарий симуляциясын біріктіреді. Модель нақты нысаннан алшақтап кетпеуі үшін калибрлеу және uncertainty бағасы қажет.

Қарапайым мысал. Зауыт конвейерінің цифрлық егізінде қозғалтқыш температурасы мен дірілі бақыланады. Жүйе техникалық қызметті бүгін жасау мен бір апта кейін жасау салдарын өндірісті тоқтатпай салыстырады.

Негізгі ой: цифрлық егіз қауіпсіз тәжірибе жасауға мүмкіндік береді, бірақ оның шешімі нақты дерекпен тұрақты жаңартылғанда ғана пайдалы.

Түсіндірілетін AI

Енді мәселені жүйелік және жауапты қолдану тұрғысынан тереңдетемін.

Түсіндірілетін AI модельдің неліктен белгілі нәтиже бергенін адамға түсінуге көмектеседі. Бұл сенім үшін ғана емес, қатені табу, әділетсіздікті байқау және шешімді қорғау үшін қажет.

SHAP тәсілінде белгі үлесі φ_i барлық мүмкін белгі жиындарындағы шекті әсерді жинақтайды. Қарапайым түсінік: нәтиже бастапқы орташа мәннен әр белгі қосқан оң немесе теріс үлес арқылы өзгереді. Түсіндірме модельдің өзі емес, оны тексеру құралы екенін ұмытпаймыз.

Қарапайым мысал. Несие тәуекелі жоғары шыққанда жүйе «кіріс аз» деп қана қоймай, төлем тарихы, қарыз жүктемесі және табыс тұрақтылығы үлесін көрсетеді. Сарапшы себептің қисынды екенін тексереді.

Негізгі ой: түсіндірме шешімді автоматты түрде дұрыс етпейді, бірақ оны талдауға және жауапкершілікпен қолдануға мүмкіндік береді.

$$f(x) = E[f(x)] + \Sigma_i \varphi_i$$

Слайд 16. Жауапты AI және әділдік

Енді алдыңғы ойды келесі технологиялық бағытпен байланыстырамын.

Жауапты AI заңдылық, қауіпсіздік, құпиялық, әділдік, түсіндірмелілік және адам бақылауын жүйенің бүкіл өмірлік циклін енгізеді. Бұл соңында қосылатын құжат емес, бастапқы талап.

Топтық оң нәтиже айырмасын $\Delta_DP=|P(\hat{y}=1|A=0)-P(\hat{y}=1|A=1)|$ арқылы бақылауға болады. Бірақ бір ғана метрика барлық әділдік мәселесін шешпейді. Контекст, қате түрі және нақты зиян бірге талданады.

Қарапайым мысал. Қызметкерді іріктеу моделінің жалпы дәлдігі жоғары болуы мүмкін, бірақ бір топтың білікті үміткерлерін жиі өткізіп алса, жүйе әділетсіз нәтиже береді. Сондықтан метрика топтар бойынша бөлініп тексеріледі.

Негізгі ой: жауапты AI орташа сапамен шектелмей, жүйенің кімге қалай әсер ететінін тексереді.

$$\Delta DP = |P(\hat{y}=1|A=0) - P(\hat{y}=1|A=1)|$$

Аудиторияға сұрақ: жалпы дәлдік жоғары болса да, жүйе әділетсіз болуы мүмкін бе?

Құпиялық және федеративті оқыту

Осы түсініктен кейін практикалық механизмге көшемін.

Дерек құндылығы артқан сайын құпиялықты сақтау маңызды болады. Data minimization тек қажетті деректі жинауды, purpose limitation оны анықталған мақсатпен пайдалануды білдіреді.

Федеративті оқытуда дерек құрылғыдан шықпайды, жергілікті модель жаңартулары біріктіріледі: $w_\{t+1\}=\Sigma_k(n_k/N)w_k$. Бұл құпиялықты күшейтеді, бірақ жаңартудан ақпарат ағуы мүмкін болғандықтан secure aggregation және differential privacy қажет.

Қарапайым мысал. Бірнеше аурухана пациент жазбасын ортақ серверге жібермей, жергілікті модель үйретеді. Орталық тек қорғалған жаңартуларды біріктіреді, ал әр ұйым рұқсат пен сапаны сақтайды.

Негізгі ой: федеративті оқыту деректі орталықтандырмай бірлесіп үйренуге көмектеседі, бірақ ол құпиялықтың жалғыз кепілі емес.

$$w_{t+1} = \Sigma_k \, (n_k/N) w_k$$

AI киберқауіпсіздігі

Бұл үрдіс алдыңғы мүмкіндікті толықтырып, жаңа талап қояды.

AI жүйесі дәстүрлі киберқауіппен қатар adversarial input, data poisoning, model theft және prompt injection сияқты жаңа тәуекелге кездеседі. Қауіпсіздік дерек, модель, API және пайдаланушы әрекетін толық қамтиды.

Zero Trust қағидасы әр сұранысты тексеруді, ең аз рұқсатты және тұрақты мониторингті талап етеді. Генеративті жүйеде сыртқы мәтін нұсқау ретінде қабылданбауы үшін дерек пен командаңың шекарасы ажыратылады.

Қарапайым мысал. RAG жүйесіне енгізілген құжатта «алдыңғы ережені елеме» деген жасырын мәтін болуы мүмкін. Қауіпсіз жүйе құжатты сенімсіз дерек деп қарап, құрал шақыру құқығын бөлек тексереді.

Негізгі ой: AI қауіпсіздігі модельді ғана емес, бүкіл ақпараттық тізбекті және рұқсат саясатын қорғайды.

Білім графтары және нейросимволдық тәсіл

Енді мәселені жүйелік және жауапты қолдану тұрғысынан тереңдетемін.

Білім графы нысандарды түйін, ал олардың байланысын қыр ретінде көрсетеді. Ол ұйымдағы термин, объект және ережені ортақ мағынада байланыстырады.

Нейросимволдық тәсіл нейрондық модельдің тану қабілетін символдық ереже мен логикалық тексерумен біріктіреді. Бұл әсіресе дәлел, шектеу және түсіндіру қажет салада пайдалы.

Қарапайым мысал. Медициналық жүйе мәтіннен «дәрі–ауру–қарсы көрсетілім» байланысын табады. Білім графы дәрілердің үйлесімін тексереді, ал ереже қауіпті комбинацияға автоматты ескерту береді.

Негізгі ой: білім графы мен ереже модель жауабына құрылым, мағына және тексеру қабатын қосады.

Адам мен AI бірлескен жұмысы

Енді алдыңғы ойды келесі технологиялық бағытпен байланыстырамын.

Заманауи үрдіс адамды толық алмастырудан гөрі адам мен AI мүмкіндігін дұрыс бөлуге бағытталады. AI үлкен деректі жылдам қарап, нұсқа ұсынады; адам мақсат, контекст, құндылық және жауапкершілікті сақтайды.

Human-in-the-loop адам әр маңызды шешімді тексереді, human-on-the-loop жүйені бақылап қажет кезде араласады, human-out-of-the-loop тек төмен тәуекелді автоматтандыруға жарайды. Интерфейс сенімділік пен белгісіздікті жасырып қалмауы керек.

Қарапайым мысал. Рентген жүйесі күмәнді аймақты белгілейді және ұқсас жағдайларды көрсетеді. Диагнозды дәрігер қояды, ал қате белгі кейін модельді жақсарту үшін кері байланысқа айналады.

Негізгі ой: тиімді бірлескен жүйе адамның рөлін анық көрсетіп, модель ұсынысын тексеруге ыңғайлы етеді.

Green AI және энергия тиімділігі

Осы түсініктен кейін практикалық механизмге көшемін.

Үлкен модельдер мен дата-орталықтар энергия, жабдық және салқындату ресурсын пайдаланады. Сондықтан Green AI тек дәлдікті емес, есептеу құнын және қоршаған орта әсерін де бағалайды.

Шартты энергияны $E=P\times t$, ал көміртек ізін $C=E\times\gamma$ арқылы есептеуге болады. Мұнда P қуат, t уақыт, γ электрдің көміртек қарқындылығы. Model compression, тиімді аппарат және жүктемені қолайлы уақытқа ауыстыру әсерді азайтады.

Қарапайым мысал. Екі модельдің дәлдігі 92% және 92,3%, бірақ екіншісі төрт есе көп энергия жұмсайды. Нақты міндетте 0,3 пайыздық пункт осы шығынға тұра ма деген сұрақ қойылады.

Негізгі ой: ең үлкен модель әрқашан ең тиімді жүйе емес; сапа, кідіріс, энергия және құн бірге өлшенеді.

$$\mathbf{E = P \cdot t, \quad C = E \cdot \gamma}$$



2-сурет. Модель сапасы мен энергия шығынының шартты байланысы

Аудиторияға сұрақ: сапаның шағын өсімі энергия шығынын қашан ақтайды?

Жекелендіру және ұсыныс жүйелері

Бұл үрдіс алдыңғы мүмкіндікті толықтырып, жаңа талап қояды.

Жекелендіру жүйесі пайдаланушыға оның мақсаты мен контекстіне сәйкес мазмұн, тауар немесе қызмет ұсынады. Қазіргі үрдіс тек өткен әрекетті қайталаудан ұзақ мерзімді пайда мен әртүрлілікті ескеруге өтуде.

Матрицалық факторизацияда болжам r_hat_ui=p_u^Tq_i түрінде есептеледі. Бірақ дәлдікпен бірге novelty, diversity және filter bubble тәуекелі қаралады. Пайдаланушыға ұсынысты өзгерту және өшіру құқығы берілуі тиіс.

Қарапайым мысал. Оқу платформасы студентке тек оңай тапсырмаларды ұсынса, қысқа мерзімде ұпай өседі, бірақ білім дамымайды. Жақсы жүйе меңгерілген және жаңа тақырыптың теңгерімін сақтайды.

Негізгі ой: жекелендіру адамның таңдауын тарылтпай, оның нақты мақсаты мен дамуына қызмет етуі керек.

$$\hat{r}_{ui} = p_u^T q_i$$

Интероперабельдік және ашық стандарттар

Енді мәселені жүйелік және жауапты қолдану тұрғысынан тереңдетемін.

Интеллектуалды жүйе дерекқор, бизнес-қосымша, сенсор, модель және адам интерфейсімен байланысады. Интероперабельдік осы бөліктердің ортақ формат пен келісілген мағынада жұмыс істеуін қамтамасыз етеді.

API, оқиға ағыны, модель форматы, metadata және ontology жүйені ауыстыру мен кеңейтуді жеңілдетеді. Ашық стандарт vendor lock-in тәуекелін азайтады, бірақ қауіпсіздік пен нұсқа үйлесімділігі бәрібір басқарылуы керек.

Қарапайым мысал. Қала платформасы көлік, ауа сапасы және төтенше қызмет дерегін ортақ API арқылы алады. Әр жүйе уақыт, географиялық координата және объект кодын бірдей түсінгенде ғана талдау дұрыс болады.

Негізгі ой: интеллектуалды жүйенің ұзақ өмірі оның модульдері мен деректерінің ашық әрі басқарылатын байланысына тәуелді.

Реттеу және AI тәуекелдерін басқару

Енді алдыңғы ойды келесі технологиялық бағытпен байланыстырамын.

Әлемде AI реттеуі тәуекелге негізделген тәсілге қарай дамып келеді. Жүйенің міндеті, ықтимал зияны, әсер ететін адамдары және бақылау шаралары құжатталады. NIST AI RMF сияқты шеңберлер тәуекелді басқаруды Govern, Map, Measure және Manage функцияларымен байланыстырады.

Жоғары әсерлі жүйеде дерек паспорты, model card, тест нәтижесі, оқиға журналы және жауапты тұлға қажет. Еуропалық AI Act те қолдану тәуекеліне қарай әртүрлі міндет жүктейді. Ұйым үшін ең дұрыс әрекет — заң күшіне енгенін күтпей, басқаруды қазірден енгізу.

Қарапайым мысал. Университет емтихан бақылау жүйесін енгізер алдында қажеттілікті, жалған күдік ықтималдығын, ерекше қажеттілігі бар студентке әсерін, шағым және адам тексеру жолын құжаттайды.

Негізгі ой: реттеу инновацияны тоқтатуға емес, жоғары әсерлі жүйені дәлелмен, бақылаумен және жауапкершілікпен қолдануға бағытталады.

Аудиторияға сұрақ: жоғары тәуекелді AI жүйесінде қандай дәлелдер сақталуы керек?

Даму жол картасы және қорытынды

Осы түсініктен кейін практикалық механизмге көшемін.

Соңында барлық үрдісті енгізу жол картасына біріктіремін. Бірінші қадам — технологияны емес, нақты мәселе мен құндылық өлшемін анықтау. Екінші қадам — дерек пен тәуекел дайындығын бағалау. Үшінші қадам — шағын пилот жасап, базалық шешіммен салыстыру.

Одан кейін MLOps, қауіпсіздік, мониторинг және адам рөлін өндірістік архитектураға енгіземіз. Пилот сәтті болса ғана масштабтаймыз. Пайда дәлдікпен шектелмейді: уақыт, құн, сапа, әділдік, энергия және пайдаланушы сенімі бірге өлшенеді.

Қарапайым мысал. Ақылды кампус жобасы алдымен бір ғимаратта энергия болжамын сынайды. Нәтиже түсінікті болып, үнем дәлелденгеннен кейін ғана басқа ғимараттарға таратылады. Қате сценарийі мен қолмен басқару жолы сақталады.

Қорытынды ой: заманауи интеллектуалды жүйе — бұл күшті модель, сенімді дерек, ашық архитектура, адам бақылауы және үздіксіз жауапты басқарудың бірлігі.

Салыстыру кестелері

1-кесте. Есептеуді орналастыру тәсілдері

Тәсіл	Күшті жағы	Шектеуі	Мысал
Бұлт	Икемді масштаб	Желіге тәуелді	Ұсыныс сервисі
Edge	Аз кідіріс, құпиялық	Ресурс шектеулі	Өндірістік камера
Гибрид	Екі тәсілді теңестіреді	Басқаруы күрделі	Ақылды көлік

2-кесте. Генеративті жүйені толықтыру тәсілдері

Тәсіл	Не өзгереді	Артықшылық	Тәуекел
Prompt	Нұсқау	Жылдам	Тұрақсыз нәтиже
RAG	Сыртқы контекст	Өзекті дереккөз	Іздеу қатесі
Fine-tuning	Модель салмағы	Мінезді бейімдеу	Құн және ұмыту
Agent tools	Әрекет мүмкіндігі	Жұмыс үдерісі	Артық рұқсат

3-кесте. Жауапты AI бақылаулары

Тәуекел	Метрика немесе дәлел	Бақылау	Адам рөлі
Қате жауап	Дәлдік, calibration	Тест және мониторинг	Күмәнді істі тексеру
Әділетсіздік	Топтық қате айырмасы	Bias аудиті	Шағымды қарау
Құпиялық	Дерек картасы	Минимизация, шифрлау	Рұқсат беру
Қауіпсіздік	Оқиға журналы	Zero Trust, red team	Тоқтату шешімі

Шағын практикалық тапсырма

Тапсырма: университетке арналған интеллектуалды кеңесші жүйенің даму жоспарын құрыңыз. 1) бір нақты пайдаланушы мәселесін таңдаңыз; 2) cloud, edge немесе hybrid архитектурасын негіздеңіз; 3) RAG үшін үш сенімді дереккөзді атаңыз; 4) адам бекітуін қажет ететін әрекетті көрсетіңіз; 5) сапа, әділдік, қауіпсіздік және энергия бойынша кемінде төрт метрика таңдаңыз; 6) пилоттан масштабтауға дейінгі үш кезеңді жазыңыз.

Күтілетін талдау: ресми оқу ережелері RAG базасына түседі, модель жауаппен бірге дереккөз көрсетеді, жеке дерек рұқсатпен өңделеді, ал күрделі өтініш қызметкерге беріледі. MLOps құжат, модель және метрика нұсқасын сақтайды.

Бақылау сұрақтары

- Edge AI қандай жағдайда бұлттан тиімді?
- Lakehouse нені біріктіреді?
- Іргелі және генеративті модель айырмасы қандай?
- RAG жауабының сапасы қандай кезеңдерге тәуелді?
- Агенттік AI үшін рұқсат шекарасы неге қажет?
- MLOps қандай объектілердің нұсқасын сақтайды?
- SHAP түсіндірмесі нені көрсетеді?
- Федеративті оқыту құпиялықты толық кепілдей ме?
- Green AI-де қандай көрсеткіштер бірге қаралады?
- Тәуекелге негізделген AI басқаруы нені талап етеді?

Негізгі терминдер

Термин	Қысқаша анықтама
Edge AI	Модельді дерек көзіне жақын орындау
Foundation model	Көп міндетке бейімделетін алдын ала үйретілген үлкен модель
RAG	Іздеу нәтижесін генерация контекстіне қосу
Agentic AI	Жоспарлап, құрал қолданып әрекет ететін AI
MLOps	Модель өмірлік циклін инженерлік басқару
Digital twin	Нақты нысанның дерекпен жаңаратын виртуалды үлгісі
XAI	Модель шешімін түсіндіру әдістері
Federated learning	Деректі көшірмей жергілікті жаңартуларды біріктіріп оқыту
Green AI	AI сапасын ресурс және энергиямен бірге бағалау
AI governance	AI рөлі, тәуекелі және бақылауын басқару жүйесі

Ұсынылатын әдебиеттер мен дереккөздер

- Stanford Institute for Human-Centered AI. AI Index Report 2025.
- NIST. Artificial Intelligence Risk Management Framework 1.0.
- NIST. Generative Artificial Intelligence Profile NIST AI 600-1.
- OECD.AI. Artificial Intelligence Policy Observatory және OECD AI Principles.
- European Commission. Artificial Intelligence Act ресми ақпараттық материалдары.
- Russell S., Norvig P. Artificial Intelligence A Modern Approach.
- Kleppmann M. Designing Data-Intensive Applications.