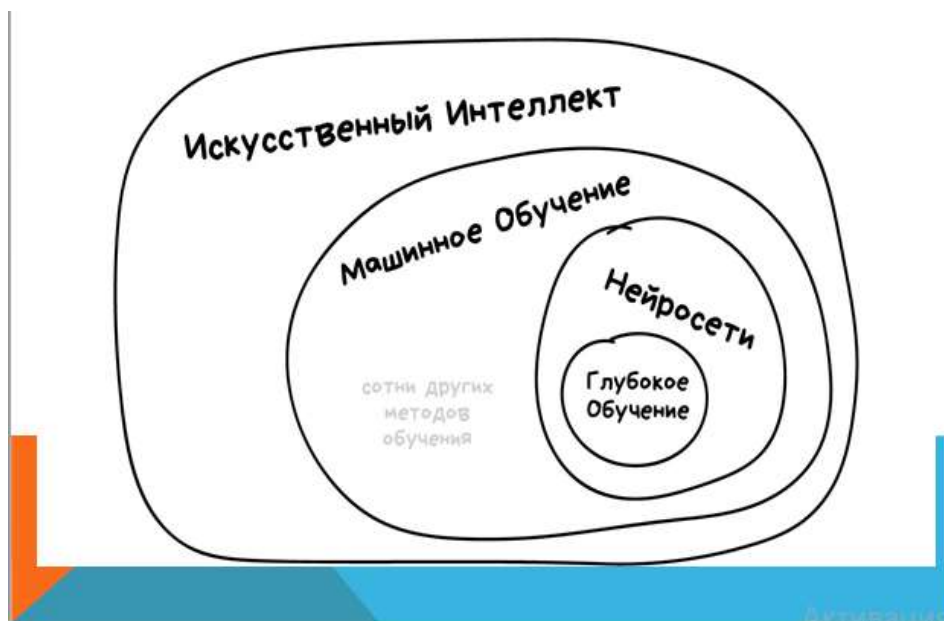


Лекция 2. Машиналық оқыту негіздері.

1. **Лекция мақсаты мен міндеттері:** Машиналық оқытудың логикалық моделі, шешімдер шежіресінің негізгі принциптері, әдістері мен міндеттері оқыту.
2. **Лекция мазмұны:**
 1. Машиналық оқытудың логикалық модельдері.
 2. Шешімдер шежіресі.
3. **Оқытудың техникалық құралдары:** Zoom онлайн оқыту платформасы
4. **Лекция оқудың тәртібі, оқыту әдістері мен түрлері:** баяндау, сұрақ – жауап, түсіндіру, кіріспе лекция

Машиналық оқыту (Machine learning, ML) – жасанды интеллект саласындағы әдістер жиынтығы, өз тәжірибесінен үйренетін машинаны жасау үшін қолданылатын алгоритмдер жиынтығы. Оқыту ретінде машина үлкен көлемді деректерді өңдейді және олардан заңдылықтарды табады. Data science және Machine learning ұғымдарын шатастырмаңыз. Бұл құралдар көптеген жолдармен қиылысады, бірақ олар әр түрлі және әрқайсысы өз міндеттерімен. Сондай-ақ, осы мақалада біз машиналық оқытуды, жасанды интеллект пен нейрондық желілерді бір үйіндіге қалай араластырмауға болатындығын бір рет түсінеміз.

Artificial (жасанды интеллект, АИ) – адамның ақыл-ойына ұқсас бағдарламалар жасауға көмектесетін әртүрлі технологиялық және ғылыми шешімдер мен әдістер. Жасанды интеллект көптеген құралдарды қамтиды, Data, Science Machine деректер туралы ғылым Машиналық оқыту.



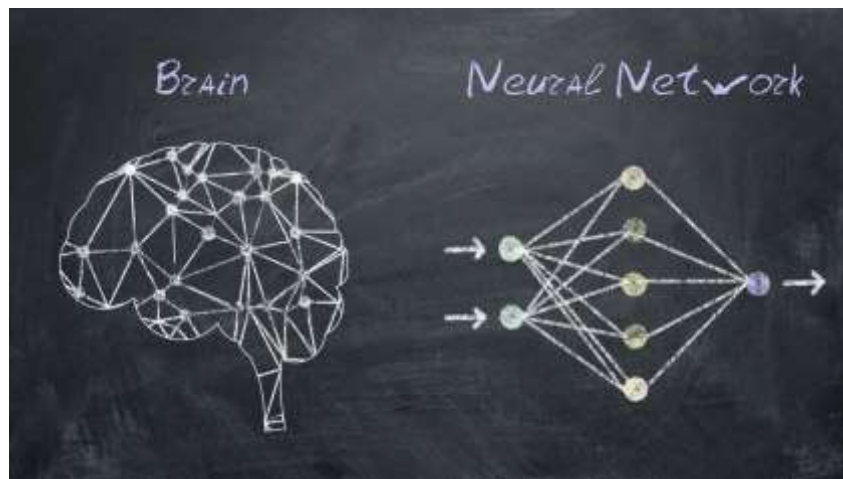
Data science-деректерді талдау және олардан құнды ақпарат, білім алу әдістері туралы ғылым. Ол машиналық оқыту және ойлау туралы ғылым (Cognitive Science) сияқты салалармен, сондай-ақ үлкен деректермен (Big Data) жұмыс істеуге арналған технологиялармен қиылысады. Data science жұмысының нәтижесі-талданған деректер және деректерді одан әрі өңдеу, сұрыптау, іріктеу, іздеу үшін дұрыс тәсілді табу.

Мысалы, кәсіпорынның қаржылық шығындары мен контрагенттердің деректері туралы келісілмеген мәліметтер бар. Бұл деректер аралық банктік деректер немесе операциялардың күні мен уақыты арқылы ғана өзара байланысты. Терең машиналық талдау нәтижесінде аралық деректер арқылы қай контрагенттің ең қымбат екенін анықтауға болады.

Машиналық оқыту немесе Machine learning – AI бөлімдерінің бірі, Алгоритмдер компьютерге қатаң белгіленген ережелерді сақтамай мәліметтер негізінде қорытынды жасауға мүмкіндік береді. Яғни, машина күрделі және көп параметрлік есептерде үлгіні таба алады (оны адам миы шеше алмайды), осылайша нақты жауаптарды таба алады. Нәтижесінде - дұрыс болжау.



Жасанды нейрондардың көмегімен нейрондық желі белгілі бір мәселені шешетін адам миының (нейрондардың) жұмысын модельдейді, алдыңғы тәжірибені ескере отырып өзін-өзі үйретеді. Әр жолы ол аз қателіктер жібереді. Нейрондық желілер-бұл жеке құрал емес, машиналық оқытудың бір түрі.



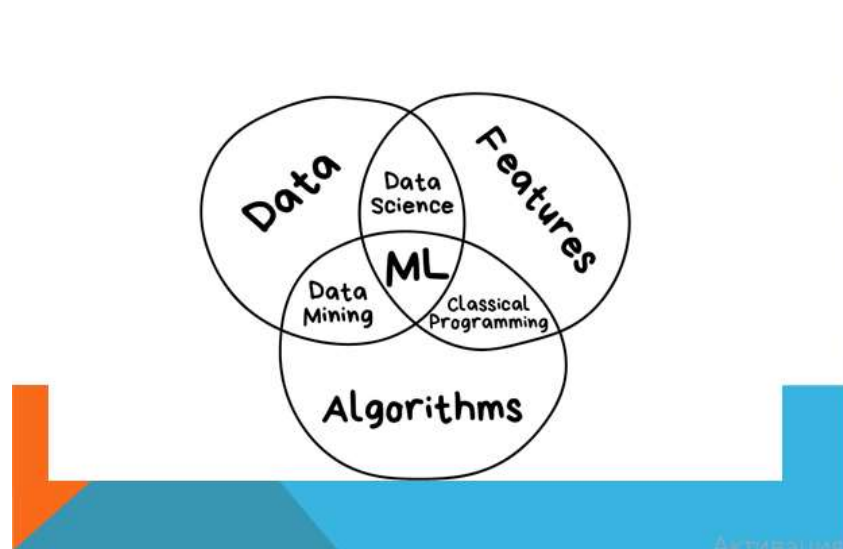
Машиналық оқытудың мақсаты және оны қолдану аясы

Машиналық оқытудың мақсаты әртүрлі күрделі аналитикалық мәселелерді шешуді ішінара немесе тіпті толық автоматтандыру болып табылады.

Сондықтан, ең алдымен, Машиналық оқыту бизнес иелері, Маркетологтар мен қызметкерлер өз жұмысында дұрыс шешім қабылдауы үшін кіріспе мәліметтер негізінде ең дәл болжам жасауға арналған. Оқыту нәтижесінде машина нәтижені болжай алады, оны есте сақтайды, қажет болған жағдайда оны ойнатады, бірнеше нұсқаның ең жақсысын тандай алады.

Қазіргі уақытта Машиналық оқыту банктерден, мейрамханалардан, жанармай құю бекеттерінен бастап өндірістегі роботтарға дейін көптеген қосымшаларды қамтиды. Күн сайын дерлік пайда болатын жаңа міндеттер машиналық оқытудың жаңа бағыттарының пайда болуына әкеледі.

Сапалы Машиналық оқыту үшін не қажет
Машиналық оқыту үш китке негізделген:



А) деректер - біз әдетте клиенттен сұрайтын негізгі ақпарат. Бұған жүйені оқыту қажет кез-келген мәліметтер жиынтығы кіреді;

Б) белгілері – жұмыстың бұл бөлігі клиентпен тығыз ынтымақтастықта жүзеге асырылады. Біз бизнестің негізгі қажеттіліктерін анықтаймыз және оқыту нәтижесінде жүйенің қандай сипаттамалары мен қасиеттерін бақылау керектігін бірлесіп шешеміз;

В) алгоритм-қойылған бизнес-міндеттерді шешу үшін әдісті таңдау. Бұл міндетті біз клиенттің қатысуынсыз, қызметкерлеріміздің күшімен шешеміз.

Деректер (Data)

Жүйеге неғұрлым көп деректерді жүктейтін болсақ, ол соғұрлым жақсы және дәл жұмыс істейді. Өздері деректер тікелей тәуелді міндеттері алдымызда машинамен.

Мысалы, поштаны маңызды хаттардан спамды сүзуге үйрету үшін мысалдар қажет. Олардың үлгісі неғұрлым көп болса, соғұрлым жақсы. Осылайша, жүйе спамның белгілері ретінде белгілі бір сөздерді – "сатып алу", "қосымша табыс", "үйде ақша табу", "ақша", "несие", "әлеуетті арттыру" – қабылдауды және осындай хаттарды жеке қалтаға жіберуді үйренеді.

Басқа тапсырмалар үшін бастапқы деректер басқаша болады. Сатып алушыға оны қызықтыруы мүмкін тауарларға кеңес беру үшін сізге сатып алу тарихы қажет. Нарықтағы бағаның өзгеруін болжау үшін баға тарихы қажет.

Жұмыстың ең қиын және сонымен бірге көлемді бөлігі-осы мәліметтерді жинау. Деректерді жинаудың екі әдісі бар: қолмен және автоматты түрде. Қол әдісі әлдеқайда баяу,бірақ дәл. Автоматты әлдеқайда жылдам, бірақ сонымен бірге көп қателіктер жібереді.

Деректердің жақсы үлгісі қымбатқа түседі, өйткені ол сіз болжаудың дәлдігіне жауап береді. Деректерді жинауды адамның ойлауымен шектемеу өте маңызды, бірақ шашыраңқы ақпаратты барынша қамтамасыз ету өте маңызды, өйткені машина адам байқамайтын жерде пайда мен өзара байланысты көре алады.

Белгілері (қасиеттері, өлшемдері, ерекшеліктері, ерекшеліктері, features).

Мысалы, автомобиль жағдайында белгілер жүгіріс, цилиндрлер саны, мүмкін болатын ең жоғары жылдамдық болады. Сатып алушы жағдайында: Жасы, жынысы, білімі, табыс деңгейі және т.б. Жануарлар жағдайында: тұқымы, бойы, құйрықтың ұшынан мұрынға дейінгі ұзындығы, түсі.

Қасиеттердің дұрыстығы сіз алған нәтижеге тікелей әсер ететіндіктен, оларды таңдау көбінесе Машиналық оқыту процесіне қарағанда көп уақытты алады. Мұнда бастысы-машиналық қабылдауды бұрмаламау үшін жеке пікірге негізделген сипаттамалар жиынтығын шектемеу. Онымен бірге соңғы нәтиже.

Белгілері бойынша құрастырылған датасет мысалы:
<https://github.com/php-ai/php-ml/blob/master/data/wine.csv>

Бұл деректер Италияның бір аймағында өсірілген, бірақ үш түрлі сорттан алынған шараптарды химиялық талдау нәтижелері болып табылады. Талдау шараптың үш түрінің әрқайсысында кездесетін 13 компонентті анықтады. Шараптың компоненттеріне сүйене отырып, оның класын анықтауға болады.

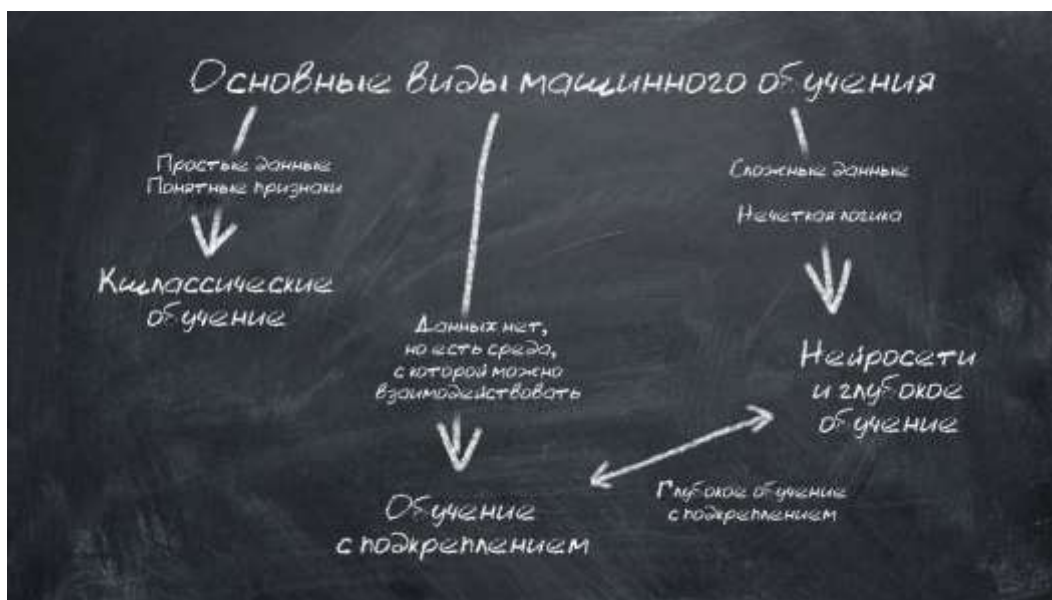
Алгоритм (Algorithm)

Белгілі бір мәселені шешу үшін жүйелі операциялар жүйесі. Басқаша айтқанда-шешім әдісі. Әрбір нақты тапсырма үшін жеке талғампаз алгоритмді таңдауға болады. Бастапқы деректерді өңдеу нәтижесінің жылдамдығы мен дәлдігі таңдалған әдіске тікелей байланысты.

Тіпті тамаша жазылған Алгоритмдер де қойылған бизнес-міндеттерді шешуге көмектеспейтін жағдайлар да болады. Мысалы, егер сіз сайттағы кросс-сатылымдардың санын көбейтуді қаласаңыз және бұл үшін сізге өнімді ұсыну алгоритмін жақсарту қажет екеніне сенімді болсаңыз. Бірақ сіздің клиенттеріңіз іздеуден тікелей сілтемелер арқылы келетінін және сайтта көрсетілген басқа өнімдерді сатып алу бойынша кеңестерді елемейтінін білмейсіз. Сондықтан, жұмысқа кіріспес бұрын, біз клиенттің проблемасының нақты себебін анықтаймыз. Егер ол техникалық болса, оны шешуге қуана көмектеседі.

Машиналық оқытудың негізгі түрлері

Бұл мақалада біз машиналық оқытудың негізгі түрлерін қысқаша қарастырамыз, болашақта олардың әрқайсысына жеке егжей-тегжейлі материал арнаймыз.



Мұғалімнің болу белгісі бойынша оқыту мұғаліммен (Supervised Learning), мұғалімсіз (Unsupervised Learning) және арматурамен (Reinforcement Learning) оқуға бөлінеді.

- Мұғаліммен бірге оқыту машинаны заттарды немесе сигналдарды тануға үйрету қажет болған кезде қолданылады. Мұғаліммен бірге оқытудың жалпы қағидасы - "міне, есік, бұл да есік, және бұл да есік".

- Мұғалімсіз оқыту "бұл нәрсе басқалармен бірдей"принципін қолданады. Алгоритмдер ұқсастықтарды зерттейді және айырмашылықты анықтап, аномалияны анықтай алады, Ерекше немесе ұқсас емес нәрсені таниды.

- Арматуралық оқыту машинаның алдында тұрған жерде қолданылады-іс-әрекеттің көптеген нұсқалары бар сыртқы ортада тапсырмаларды дұрыс орындау. Мысалы, компьютерлік ойындарда, сауда операцияларында, ұшқышсыз техникаға арналған.

Қолданылатын алгоритмдердің түріне сәйкес екі түрді бөлуге болады:

А) классикалық оқыту-негізінен 50 жылдан астам уақыт бұрын статистикалық бюролар үшін әзірленген белгілі және жақсы зерттелген оқыту алгоритмдері. Ең алдымен деректермен жұмыс істеу міндеттеріне сәйкес келеді: жіктеу, кластерлеу, регрессия және т.б. болжау, клиенттерді сегменттеу және т. б. үшін қолданылады.

Б) нейрондық желілер және терең оқыту – машиналық оқытудың ең заманауи тәсілі. Нейрондық желілер суреттер мен бейнелерді тану немесе құру, күрделі басқару немесе шешім қабылдау алгоритмдері, машиналық аударма және ұқсас күрделі тапсырмалар қажет болған жерде қолданылады. Бірнеше тәсілдерді біріктіруге болады, содан кейін сіз модельдер ансамбльдерін аласыз.

Машиналық оқытудың басты пайдасы

Шебер тәсілмен машиналық оқытудың әртүрлі түрлерін біріктіре отырып, күнделікті бизнес-процестердің көпшілігін автоматтандыруға қол жеткізуге болады. Басқаша айтқанда, Машиналық оқыту арқылы дайындалған алгоритмдер мен роботтар күнделікті жұмысты орындай алады. Барлық шығармашылық бөлігі адамдар үшін қалады: стратегияларды құру, келіссөздер жүргізу, келісімшарттар жасау және т.б. Бұл маңызды фактор, өйткені машина өзіне берілген шеңберден шыға алмайды, ал адам миы ойланбастан ойлана алады.

Машиналардың сипаттамаларын сапалы талдау сізге клиенттерді тарту үшін көп күш жұмсау керектігін айтады, ал адамдардың міндеті – бұл күш-жігердің стратегиясын ойластыру. Сіздің бизнес міндеттеріңізді шешу үшін машиналық оқыту мен жасанды интеллектті қалай жақсы пайдалану керектігін білгіңіз келе ме? Бізге хабарласыңыз, Сіздің бизнесіңізге ең қолайлы нұсқаны айтыңыз.

5. Деңгейлік тапсырмалар:

1-деңгей.

- Жасанды интеллект дегеніміз не?
- **Data science** дегеніміз не?

2-деңгей.

- Нейронды желі дегеніміз не?
- Нейронды желінің қызметі қандай?

3-деңгей.

- Машиналық оқыту мақсаты мен қолдану саласы қандай?
- Машиналық оқытудың үш құрамдас бөліктері?
- Машиналық оқытуда берілгендермен қалай жұмыс жасайды?

6. БОӨЖ тапсырмалары:

РВМ тиісті векторларының әдісі

7. БӨЖ тапсырмалары:

Ықтималдылықтың максималды әдісімен байланыс.

8. Пайдаланылған әдебиеттер:

1. Ғ. Айғараева, Қ. Асанова, М. Нысанов. Сандық қондырғылар және микропроцессорлық жүйелер. Оқулық. 2010
2. Д. С. Козен. Есептеу теориясы. Сандық. Оқулық Алматы. 2013.
3. Асамбаев А.Ж. Жасанды интеллект негіздері.- Алматы: ЖШС РПБК "Дәуір", 2011-135б. Оқу құралы.
4. Рассел С., Норвиг П. Жасанды интеллект: Жаңашыл әдіс -Алматы, 2014.. Оқу құралы.
5. К. Пол. Цифровое искусство. Учебник. Москва , 2017, 220
6. Андреас Мюллер, Сара Гвидо. Введение в машинное обучение с помощью Python. –Москва. 2016.
7. В.Н.Негода. Машинно-ориентированное программирование. Учебное пособие Ульяновск 2015.