Пәні: Деректерді интеллектуалды талдау және визуализациялау

Деректерді интеллектуалды талдау негіздері

Деректерді интеллектуалды талдау ұғымы.

Деректер жиынтығы және олардың атрибуттары

Деректерді интеллектуалды талдау және визуализациялау

Дәріс №1

Шаяхметова А.С., PhD, қауымдастырылған профессор

Жасанды интеллект және Big Data кафедрасы

2026-2027 оқу жылы

Жұмыстың мақсаты

Деректерді интеллектуалды талдаудың теориялық негіздерімен танысу. «Деректерді интеллектуалды талдау» ұғымының мәнін ашу. Деректер жиынтығының құрылымы мен түрлерін қарастыру. Деректер атрибуттарының типтерін ажыратуды үйрену.

Жұмыс жоспары

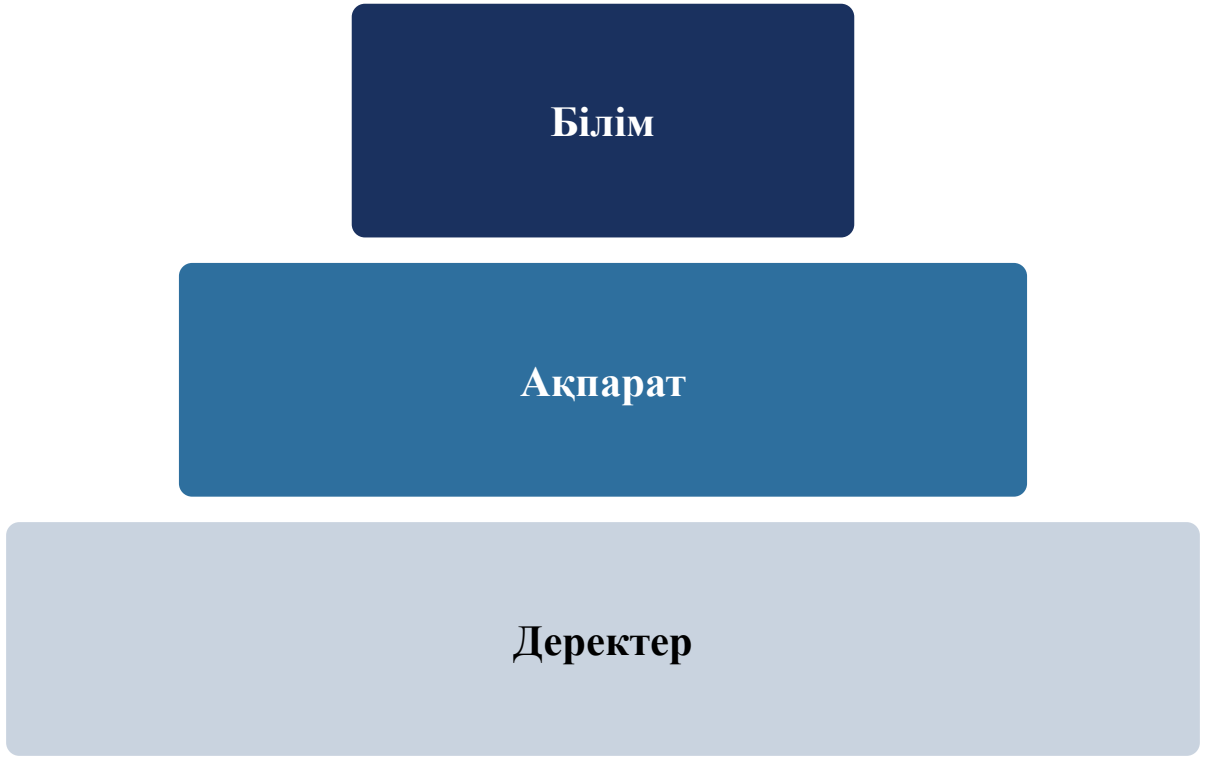
1. Деректерді интеллектуалды талдау негіздері

2. Деректерді интеллектуалды талдау ұғымы

3. Деректер жиынтығы

4. Деректер атрибуттары

Деректер, ақпарат және білім



- Деректер - өңделмеген фактілер мен өлшемдер;
- Ақпарат - мәнмәтінмен өңделген деректер;
- Білім - шешім қабылдауға жарамды заңдылықтар.
- Интеллектуалды талдау деректерден білім алуға бағытталған.

Интеллектуалды талдаудың іргелі салалары



Талдау деңгейлері

Түрі	Негізгі сұрақ	Мысал
Сипаттамалық	Не болды?	Айлық сатылым есебі
Диагностикалық	Неге болды?	Сатылым төмендеуінің себебі
Болжамдық	Не болады?	Келесі айдың сұранысын болжау
Нұсқаулық	Не істеу керек?	Тиімді бағаны ұсыну

Деректерді интеллектуалды талдау анықтамасы

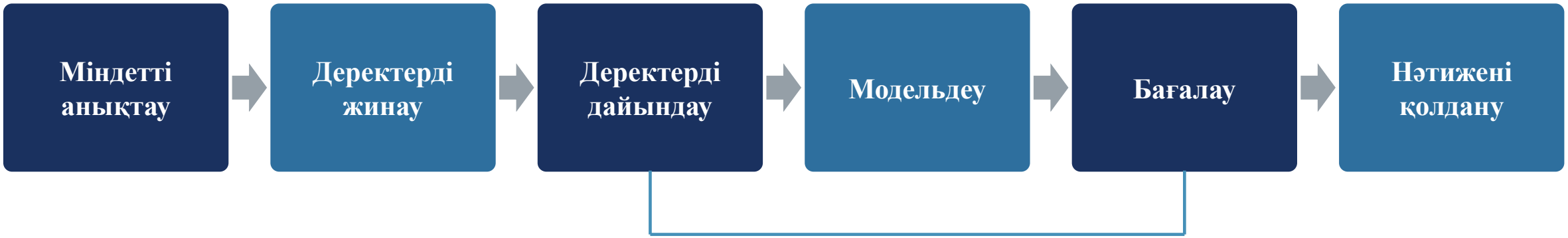
Деректерді интеллектуалды талдау (Data Mining) - үлкен көлемді деректерден бұрын белгісіз, бірақ практикада пайдалы заңдылықтарды, байланыстар мен тенденцияларды табу процесі.

Жаңалық
бұрын белгісіз білім алу

Автоматтандыру
алгоритмдер мен модельдер

Пайдалылық
шешім қабылдауға көмек

Талдау процесінің кезеңдері



Процесс қайталанбалы: нәтиже сапасы төмен болса, алдыңғы кезеңдерге оралып, деректер мен модельді жетілдіреді.

Интеллектуалды талдаудың негізгі міндеттері

Классификация
объектіні алдын ала белгілі сыныпқа жатқызу

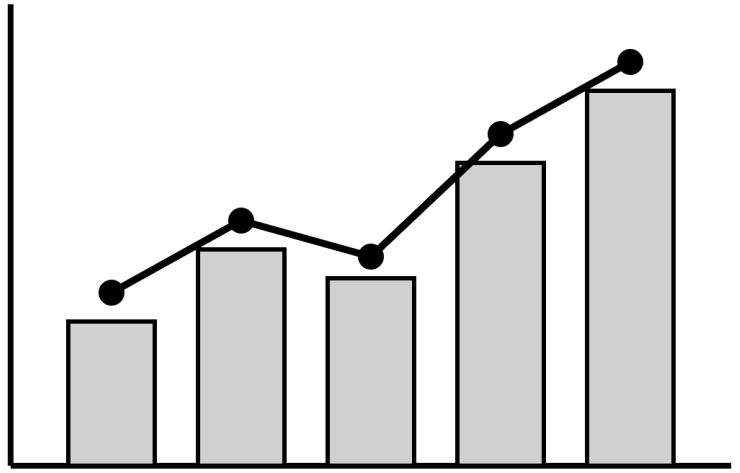
Кластерлеу
ұқсас объектілерді топтарға біріктіру

Регрессия
сандық мәнді болжау

Ассоциациялық ережелер
бірге кездесетін құбылыстарды табу

Қолдану салалары

Сала	Қолдану мысалы
Медицина	Ауруды ерте диагностикалау
Банк ісі	Несиелік тәуекелді бағалау
Сауда	Тұтынушы себетін талдау
Білім беру	Үлгерімді болжау
Өндіріс	Жабдық ақауын болжау



Деректер жиынтығы ұғымы

Деректер жиынтығы (dataset) - талдауға арналған объектілер туралы жүйеленген деректер жиыны. Әр объект бірнеше атрибут арқылы сипатталады.

Объект
жиынтықтағы жеке жазба: жол, үлгі, жағдай

Атрибут
объектінің сипаттамасы: баған, өріс, айнымалы

Мән
атрибуттың объектідегі нақты көрінісі

Деректер жиынтығының құрылымы

Атрибуттар (бағандар)

Объектілер (жолдар)

№	Жасы	Қала	Сома, теңге
1	25	Алматы	12 000
2	41	Астана	8 500
3	33	Шымкент	15 200

Қою жиекпен — мән (мысалы, «Астана»). Шартты мысал.

Деректер жиынтықтарының түрлері

Түрі	Сипаттамасы	Мысал
Кестелік	Жолдар мен бағандар	Тұтынушылар кестесі
Уақыттық қатар	Уақыт бойынша реттелген мәндер	Күндік температура
Мәтіндік	Құрылымсыз мәтін	Пікірлер, мақалалар
Графтық	Түйіндер мен байланыстар	Әлеуметтік желі
Мультимедиялық	Сурет, аудио, бейне	Медициналық суреттер

Деректер жиынтығының сапасы

Толықтық
жетіспейтін мәндер аз болуы керек

Дұрыстық
қателер мен шулар болмауы керек

Дәйектілік
деректер бір-біріне қайшы келмеуі керек

Өзектілік
деректер жаңа әрі міндетке сәйкес болуы керек

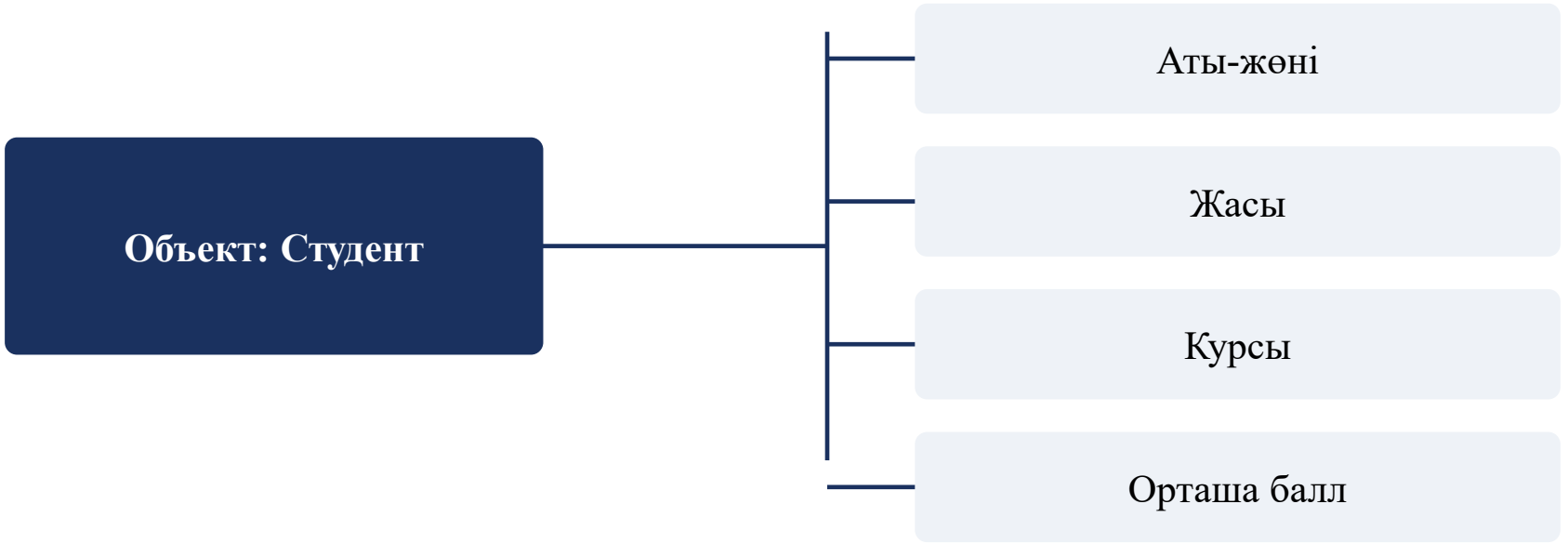
Деректерді талдауға дайындау

1. Тазалау → 2. Интеграция → 3. Түрлендіру → 4. Редукция

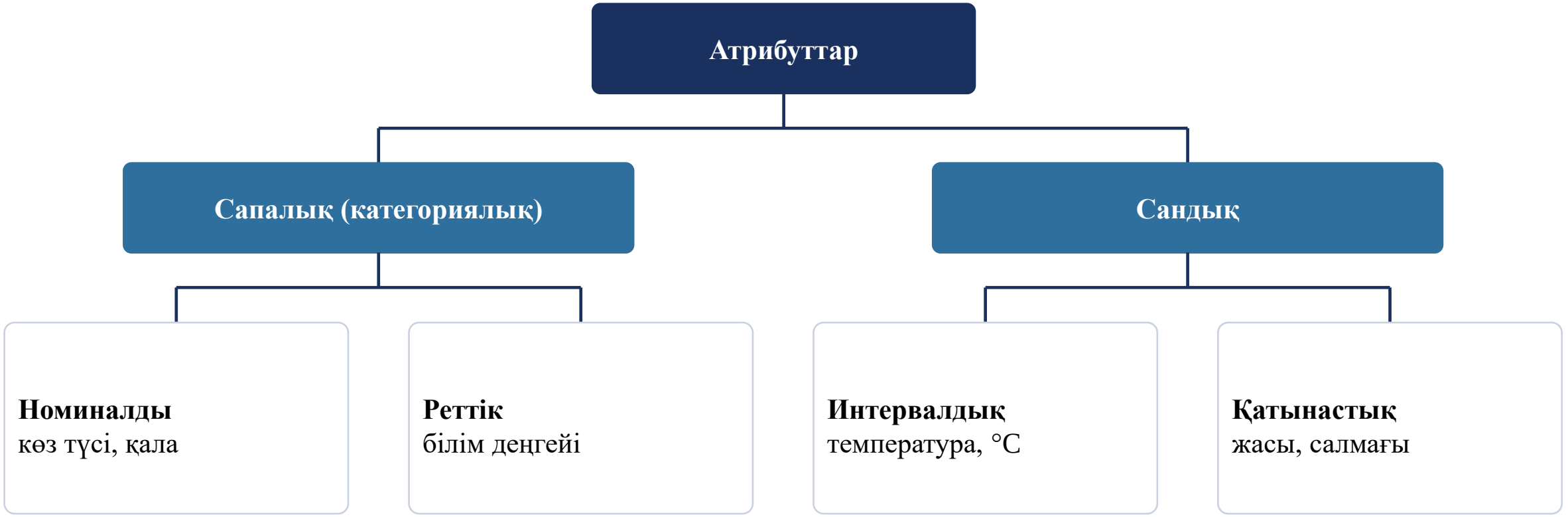
Кезең	Мазмұны
Тазалау	Жетіспейтін мәндер мен қателерді түзету
Интеграция	Әртүрлі дереккөздерді біріктіру
Түрлендіру	Нормалау, кодтау, масштабтау
Редукция	Атрибуттар мен жазбалар санын азайту

Атрибут ұғымы

Атрибут - объектінің қасиетін немесе сипаттамасын білдіретін өріс. Басқа атаулары: айнымалы, белгі (feature), өлшем, баған.



Атрибут типтерінің жіктелуі



Атрибут типтерінің сипаттамасы

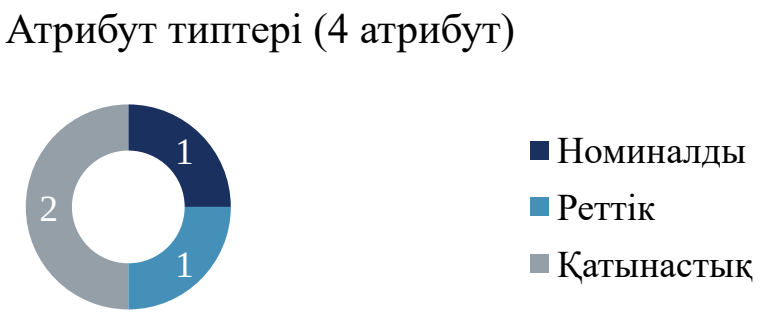
Түрі	Ерекшелігі	Мысал
Номиналды	Мәндер тек атау, реті жоқ	Қала, көз түсі
Реттік	Мәндердің реті бар, айырмасы өлшенбейді	Білім деңгейі
Интервалдық	Айырма мағыналы, нөлі шартты	Температура, °C
Қатынастық	Нөлі абсолютті, қатынас мағыналы	Жасы, салмағы, табыс

Мысал: студенттер жиынтығы

№	Мамандық	Курс	Орташа балл	Жасы
Атрибут типі	номиналды	реттік	қатынастық	қатынастық
1	Информатика	2	3,4	19
2	Экономика	3	3,1	20
3	Математика	1	3,8	18

Бір жиынтықта әртүрлі типті атрибуттар кездеседі, сондықтан талдау әдісі атрибут типін ескеріп таңдалады. Деректер шартты.





Қорытынды

1. Деректерді интеллектуалды талдау - деректерден пайдалы білім алу процесі;

2. Талдау кезеңдері: міндетті қою, дайындау, модельдеу, бағалау, қолдану;

3. Деректер жиынтығы - объектілер мен олардың атрибуттарынан тұратын құрылым;

4. Атрибут типі (номиналды, реттік, интервалдық, қатынастық) талдау әдісін таңдауды анықтайды.

Деректерді интеллектуалды талдау негіздері

Деректерді интеллектуалды талдау ұғымы.

Деректер жиынтығы және олардың атрибуттары

Дәріс №1