

Інформаційна технологія аналізу відтоку клієнтів телеком-компанії: серверна частина

Виконав: ст.гр. 2КН-17м

Уштаніт В.В.

Науковий керівник:

д.т.н., професор Яровий А.А.

Об'єкт, предмет та мета дослідження

- Об'єктом дослідження є процес аналізу відтоку клієнтів телеком-компанії.
- Предметом дослідження є програмні засоби аналізу відтоку клієнтів телеком-компанії з використанням методів машинного навчання.
- Метою комплексної кваліфікаційної роботи є підвищення точності прогнозування відтоку клієнтів телеком-компанії за допомогою використання методу випадкового лісу.

Актуальність

Сьогодні більшість продуктових компаній хочуть автоматизувати різного роду процеси в компанії, спрямовані на скорочення витрат та підвищення прибутків. Для продуктових компаній завжди були і будуть важливі клієнти, тому вони прикладають зусиль на отриманні нових та збереженні старих клієнтів, яким надаються ті чи інші послуги компанії.

Основний підхід до аналізу відтоку

1. Розуміння бізнесу
2. Розуміння даних
3. Перед-обробка даних
 1. Вибір даних
 2. Очищення даних
 3. Перетворення даних
 4. Вибір характеристик
4. Вибір методу аналізу відтоку клієнтів
5. Оцінка

Огляд моделей та методів аналізу відтоку

Проаналізовано наступні моделі та методи:

- Візуальний аналіз даних клієнтів.
- Метод дерев рішень.
- Лінійні моделі.
- Композиції методів. Метод випадкового лісу

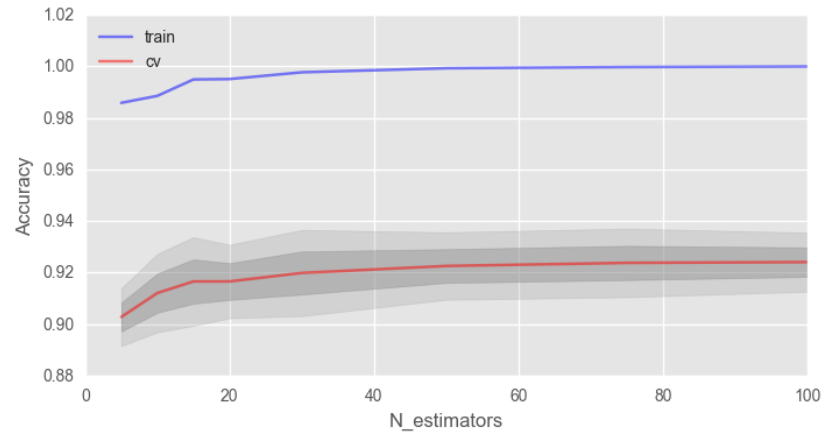
Обрано метод випадкового лісу, тому що він показав найвищий показник точності.

Алгоритм побудови випадкового лісу

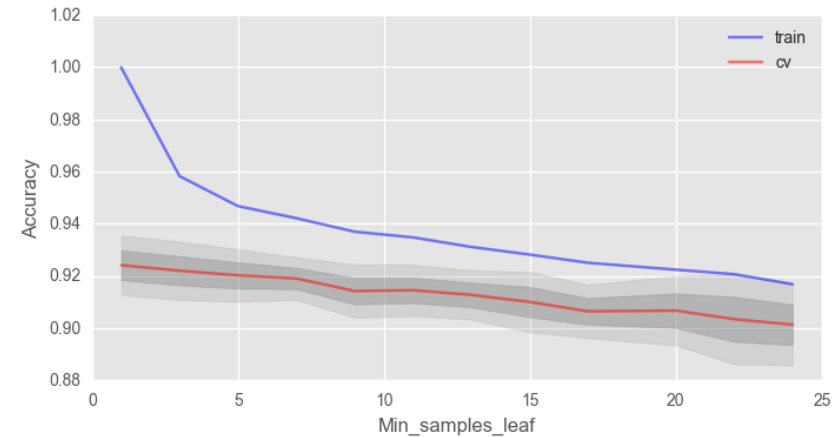
- Для кожного $n = 1, \dots, N$:
- Згенерувати вибірку X_n ;
- Побудувати вирішальне дерево b_n за вибіркою X_n :
 - по заданому критерію ми вибираємо кращий ознака, робимо розбиття в дереві по ньому і так до вичерпання вибірки;
 - дерево будується, поки в кожному аркуші не більше n_{min} об'єктів або поки не досягнемо певної висоти дерева;
 - при кожному розбитті спочатку вибирається m випадкових ознак з n вихідних і оптимальний розподіл вибірки шукається тільки серед них.
- Вибір рішення голосуванням.

Процес підбору оптимальних параметрів моделі випадкового лісу

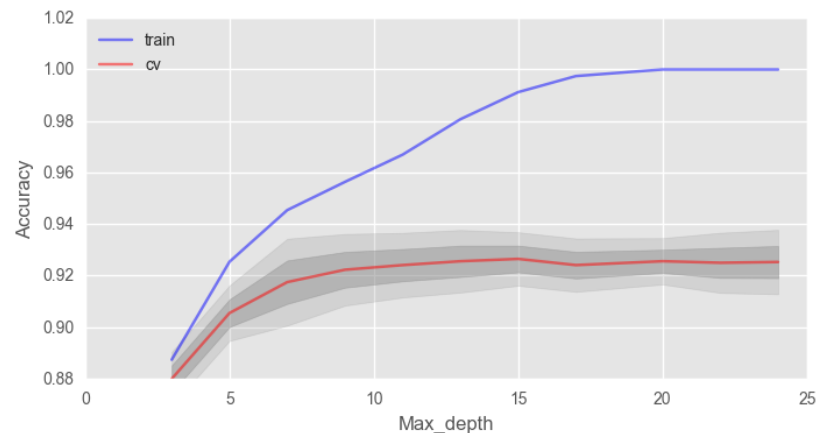
- **точність**



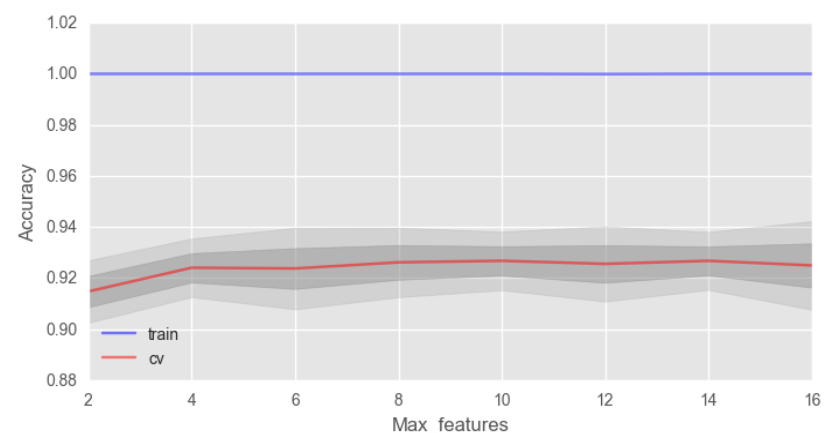
- **мінімальна кількість листків**



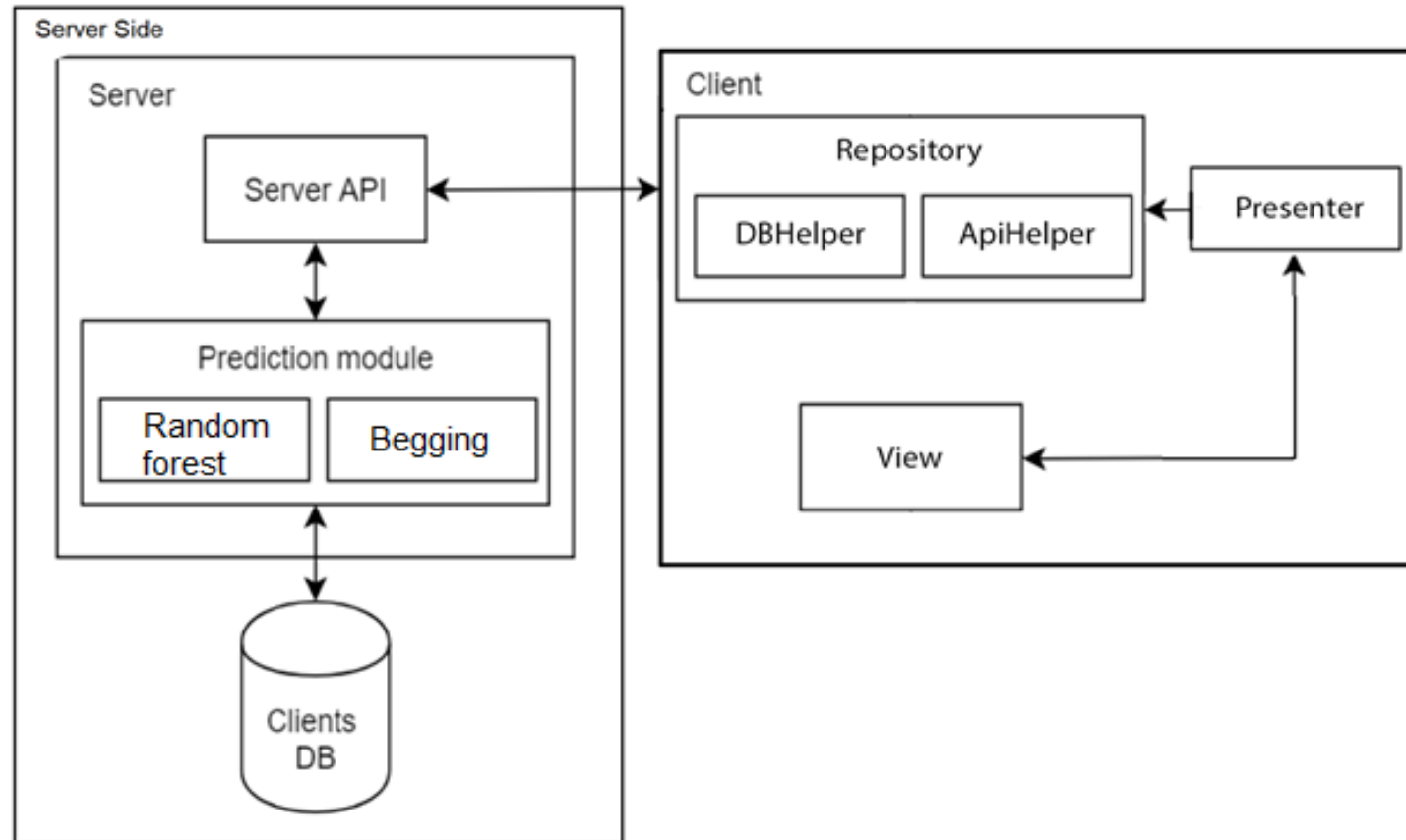
- **максимальна глибина**



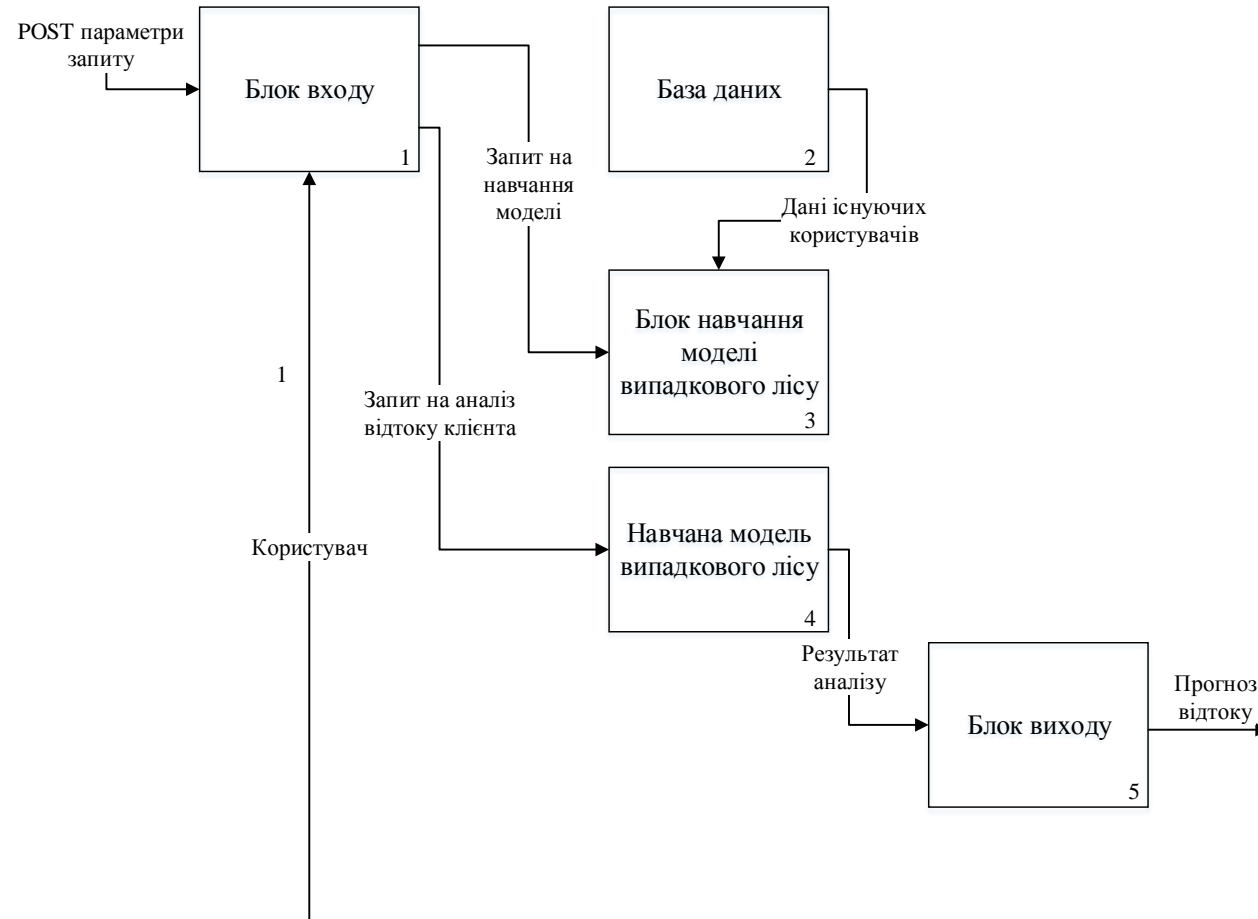
- **максимальна кількість ознак**



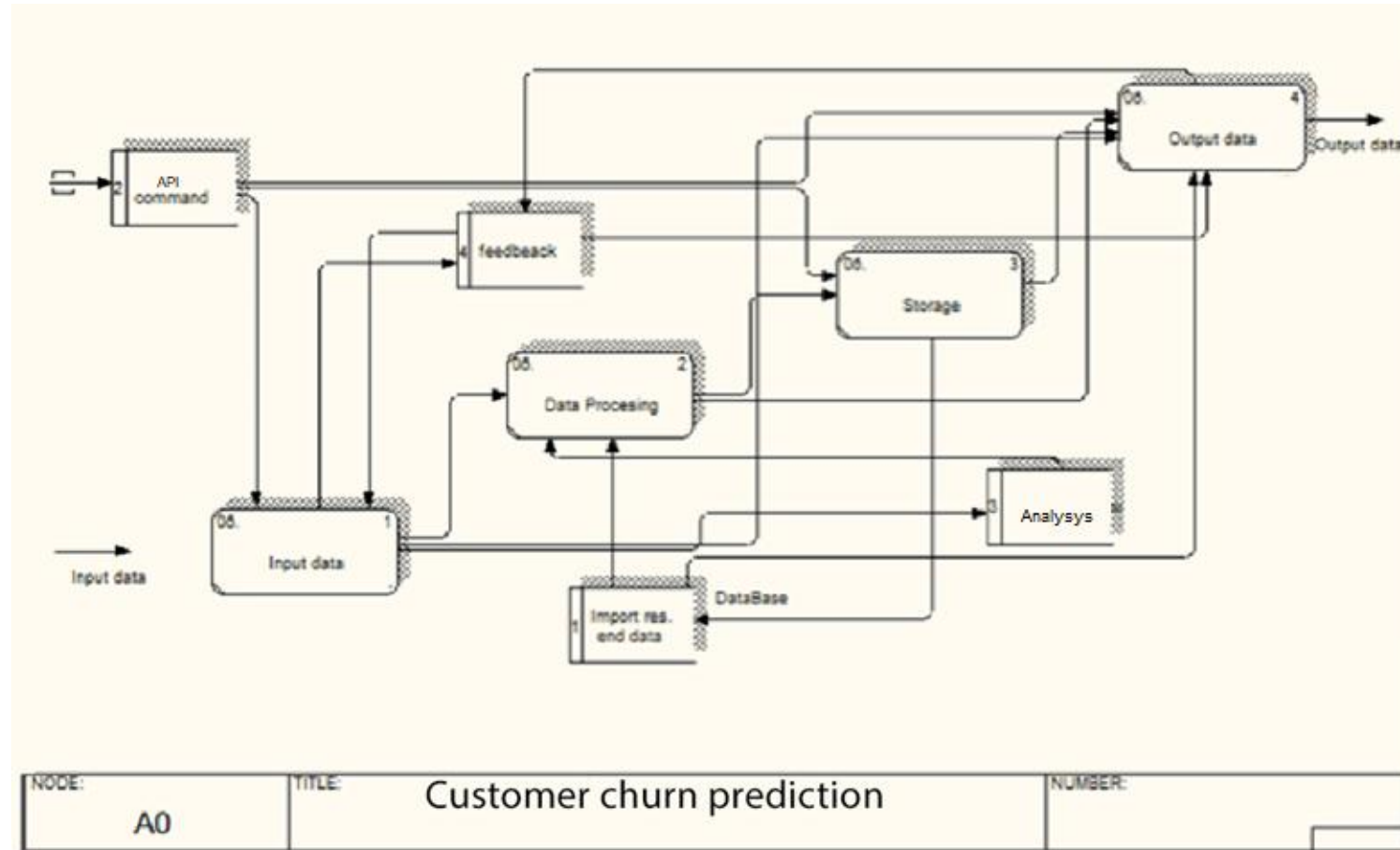
Структура інформаційної технології



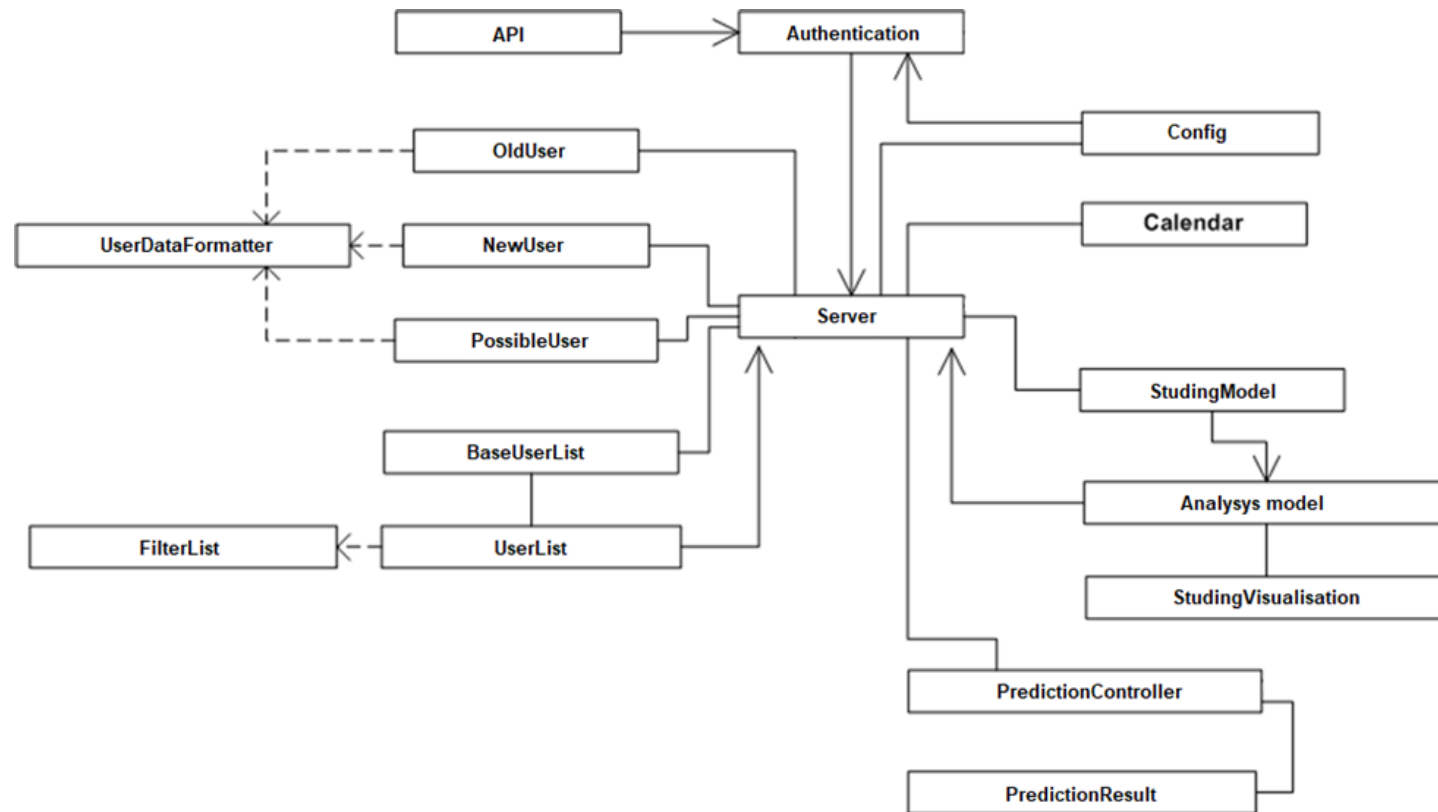
IDEF0 діаграма другого рівня декомпозиції



DFD діаграма серверної частини інформаційної технології аналізу відтоку



UML-діаграма класів серверної частини інформаційної технології аналізу відтоку



Порівняння роботи розробленого програмного забезпечення з існуючим аналогом

Під час проведення тестування було проаналізовано формування прогнозу відтоку для тестової вибірки розміром 300 клієнтів.

Параметри	Швидкість аналізу відтоку	Точність прогнозу відтоку
Програми		
Розробка	1 секунда	92%
Qlik View	2 секунди	78%

Основні результати роботи

- Вдосконалено модель аналізу відтоку клієнтів, що відрізняється від відомих моделей застосуванням методу випадкового лісу, що забезпечило підвищення точності прогнозу відтоку клієнтів телеком-компанії до 92%;
- Розроблено інформаційну технологію аналізу відтоку клієнтів, що відрізняється від відомих комбінуванням методів машинного навчання, що забезпечило підвищенню швидкодії аналізу відтоку телеком-компанії.
- Основні теоретичні та практичні результати роботи представлені на одній конференції, подано до розгляду 2 свідоцтва про реєстрацію авторського права на твір.
- Практична цінність роботи підтверджується актом впровадження її результатів.

Дякую за увагу!