

Адаптивний метод прийняття рішень для ринку цінних паперів на основі ефектуаційної концепції

Виконав:
студент 4-го курсу
Групи Ка-51
Канцедал Г. О.

Керівник:
к. ф.-м. н., доцент
Каніовська І. Ю.

Актуальність дослідження

- ▶ Задача прийняття рішень на ринках цінних паперів є провідною для сучасного світу.
- ▶ Побудова стратегій з низьким ступенем ризику є оптимальним вибором для збереження пасивів будь-якого підприємства.
- ▶ Побудована модель може бути легко перенесена на будь-який ринок, в основі якого лежить часовий ряд та контрагентами ряду виступають рівноправні користувачі без значної переваги.

► Об'єкт дослідження

Фінансові часові ряди цінних паперів, історія торговельних операцій.

► Предмет дослідження

Методи класифікації, що базуються на машинному навчанні.

► Мета дослідження

Проаналізувати предмет дослідження, реалізувати безперервний доступ до трейдингових даних, провести тестування побудованої моделі на історичних даних та в режимі реального часу.

Постановка задачі

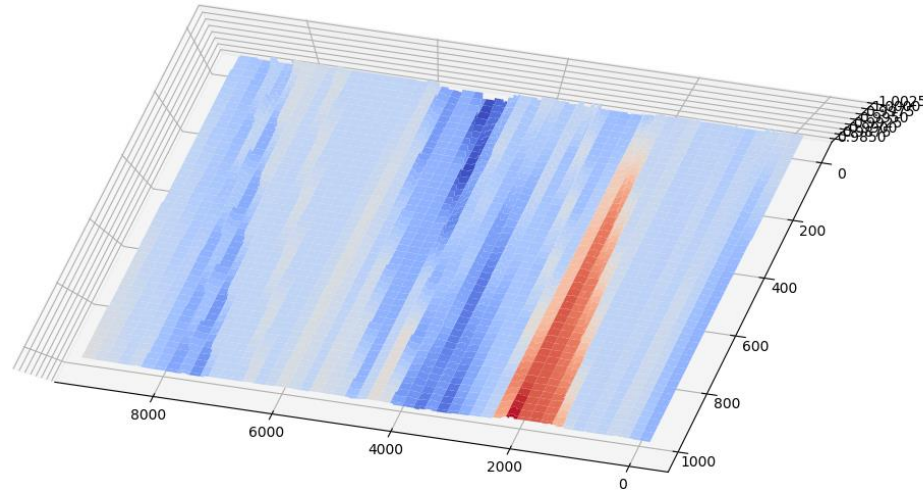
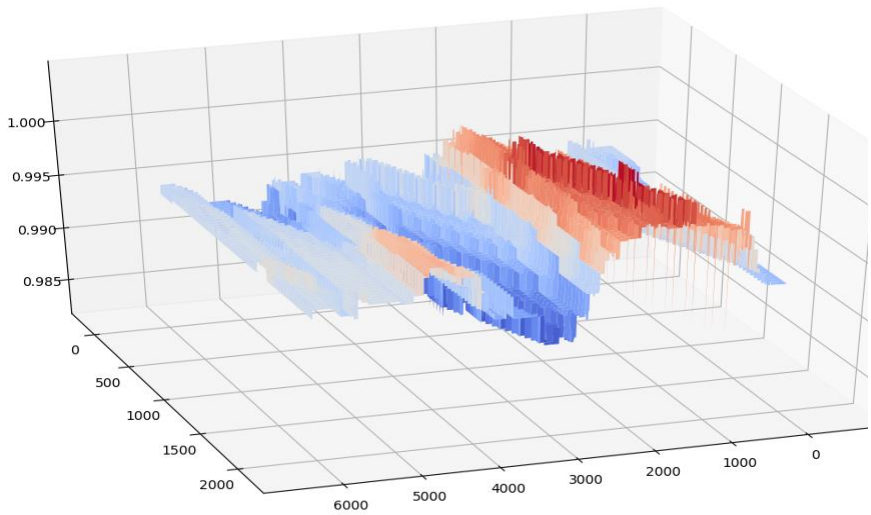
- ▶ Розробити підходи для безпосереднього доступу до даних торгової біржі, накопити історичні данні для попереднього аналізу.
- ▶ Обрати відповідну торгівельну пару для подальшої розробки та сформувавши критерій успішної роботи системи.
- ▶ Сформувавши правила, що відтворюють торгівельні операції на біржі та протестувати отриману систему в реальному часі.

Формалізація задачі

- ▶ Поставлена задача була зведена до задачі класифікації у n -вимірному просторі класу точок, що розподіляються на 3 групи - нейтральні, закупівлі, продаж.
- ▶ Побудовано моделі, що ґрунтуються на 113 вимірах (несуть інформацію за відповідний момент часу), 155 вимірах (відомості про досліджуваний момент та зміни його порівняно з попереднім) та 255 (інформація за останні 2 виміри та їх зміни).
- ▶ Подальше введення інших не дало покращення результатів.

Формалізація задачі

- ▶ В результаті формалізації отримали наступні результати (по осі $[0,6000]$ - час купівлі активу $[0,2000]$ - час утримання активу) Точки що мають темно-коричневий колір були класифіковані як час купівлі, а відповідна їм ордината як час продажу



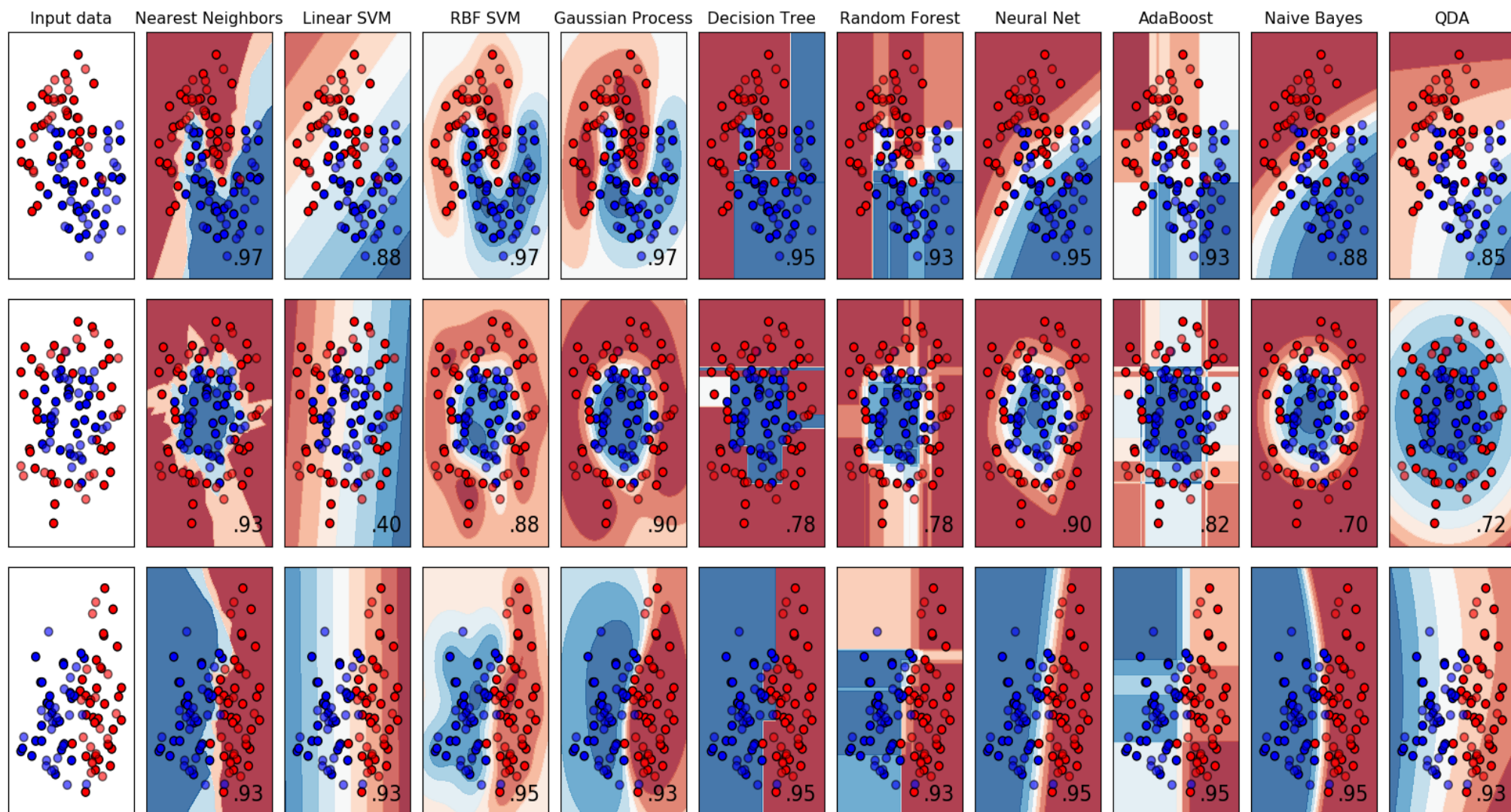
Аналіз наявних методів класифікації

- ▶ Для вирішення задачі були розглянуті наступні методи: Метод найближчих сусідів; Лінійний SVM метод; RBF SVM; Метод побудований на гаусівських процесах; Дерево прийняття рішень; Випадковий ліс (комбінація дерев прийняття рішень); Нейронні мережі; ADa boost; Баєсівські мережі; QDA.
- ▶ Найгірші результати, з позиції довгострокового прогнозування має метод AdaBoost, оскільки має білі області, що у випадку класифікації дасть велику групу точок, що не будуть віднесені до жодного з класів і відразу стануть похибкою моделі (причому частина тренувальних точок потрапила в ці області, що зовсім погано).

Аналіз наявних методів класифікації

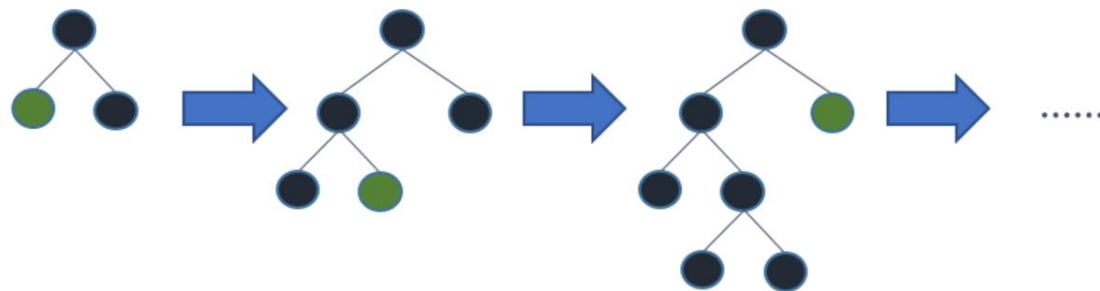
- ▶ Методи Decision tree та Random forist базуються на алгоритмі, що дозволяє розділяти данні в 2 мірному просторі лише горизонтальними та вертикальними лініями (точніше Random forist базується на Decidion tree). Однак в методі Random forist спостерігається перенавчання: велика кількість невеликих областей з високою імовірністю (з однією тренувальною точкою) та значна кількість областей з відсутньою характеристикою.
- ▶ Лінійний SVM метод показав себе повністю не пристосованим до поставленої задачі.
- ▶ Методи найближчих сусідів, RBF SVM, гаусівський процес показали непогані результати однак потребували значно більшу кількість обчислювальних ресурсів і в по загальному відхиленню не надто відрізнялись від дерев прийняття рішень.

Аналіз наявних методів класифікації

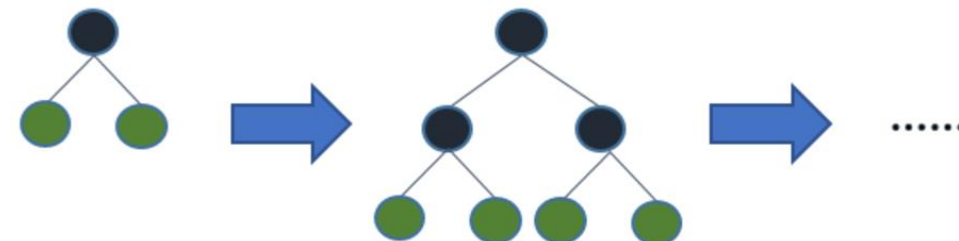


Результати аналізу

- ▶ Обрано технологію gradient boosting на базу дерев рішень (а точніше модифікацію LGBM).
- ▶ Основною відмінністю від інших алгоритмів побудованих на деревах рішень є вертикальне (leaf-wise -методологія нарощування листків і вузлів дерева, що орієнована на ефективну побудову усього дерева) нарощування листків дерева



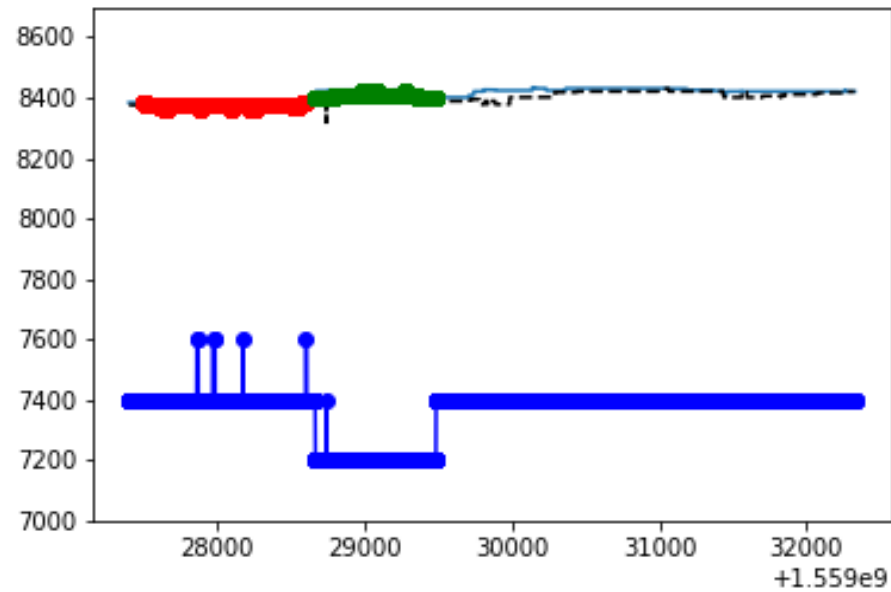
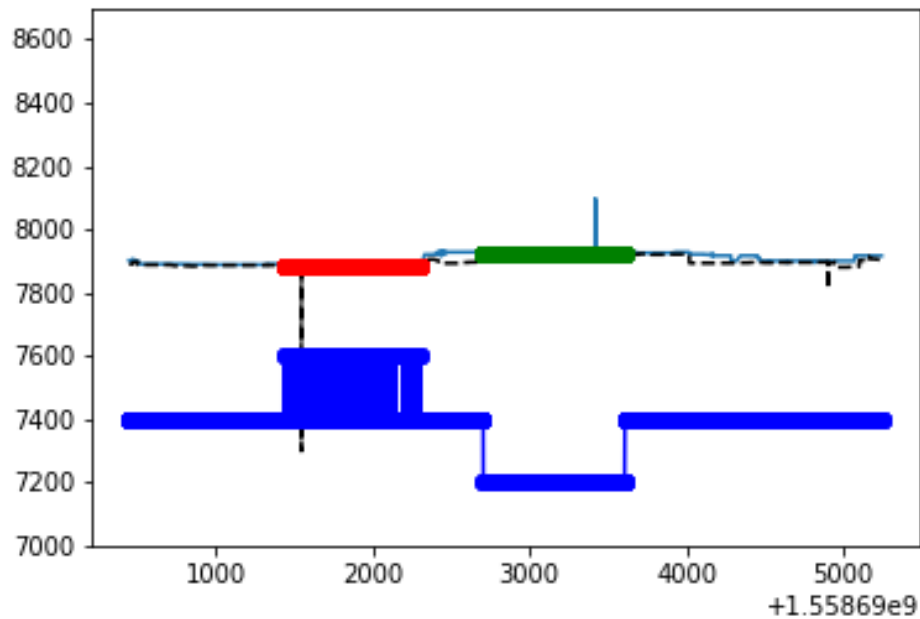
вертикальне нарощування графу



Горизонтальний ріст дерева

Робота моделі

Робота класифікатора побудованого на основі технології LGBM



Робота класифікатора на навчальних даних та тестових даних

Висновки

- ▶ Модель побудована на базі ефектуаційної концепції дозволяє отримувати прибуток з низького початкового капіталу з невисоким ступенем ризику.
- ▶ Отримані результати свідчать досягнення прибутковості моделі на рівні середньої банківської ставки : 8% в місяць.
- ▶ Досягнуто адаптивності алгоритму для будь-якого активу та можливість довгої автономної роботи за рахунок навчання моделі
- ▶ За допомогою моделі було виявлено приховані закономірності на ринку крипто валют.

Подальші дослідження

- ▶ Переглянути більшу кількість валютних пар, розглянути можливості між біржової торгівлі.
- ▶ Залучити більшу кількість параметрів до моделі
- ▶ Провести тестування моделі з більшою дискретизацією та на більших проміжках часу, виявити оптимальний час роботи моделі без навчання.

Дякую за увагу!