

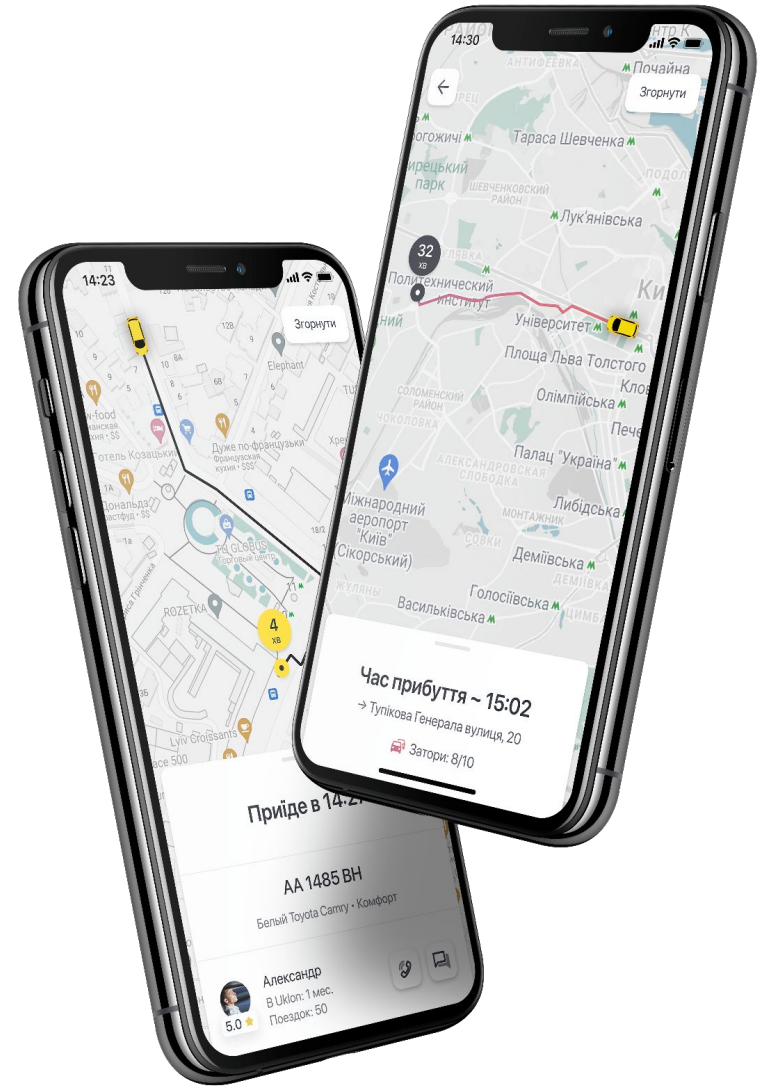
Очікуваний час прибуття (ETA)



Introduction

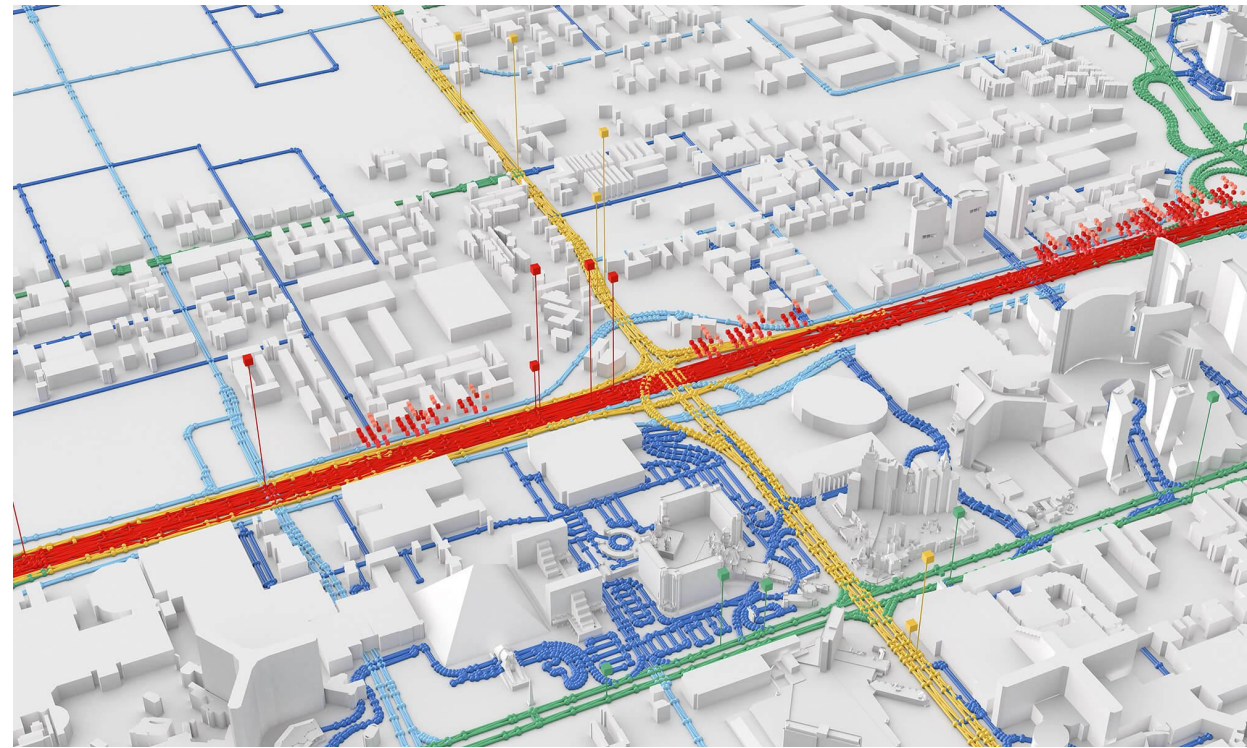
Розрахунок очікуваного часу прибуття (ETA) є однією з найважливіших послуг ride-hailing платформ.

Користувач відкриває додаток, замовляє потрібний клас авто та, після підтвердження замовлення, бачить ETA — прогнозований час, за який він прибуде до місця призначення.



Goal

Необхідно спрогнозувати час, за який користувач прибуде до місця призначення. Використовуючи для цього історичні дані поїздок з фактичним часом завершення поїздки.



Data

Зразок даних з міста Одеса.

Файл **orders.csv** містить дані:

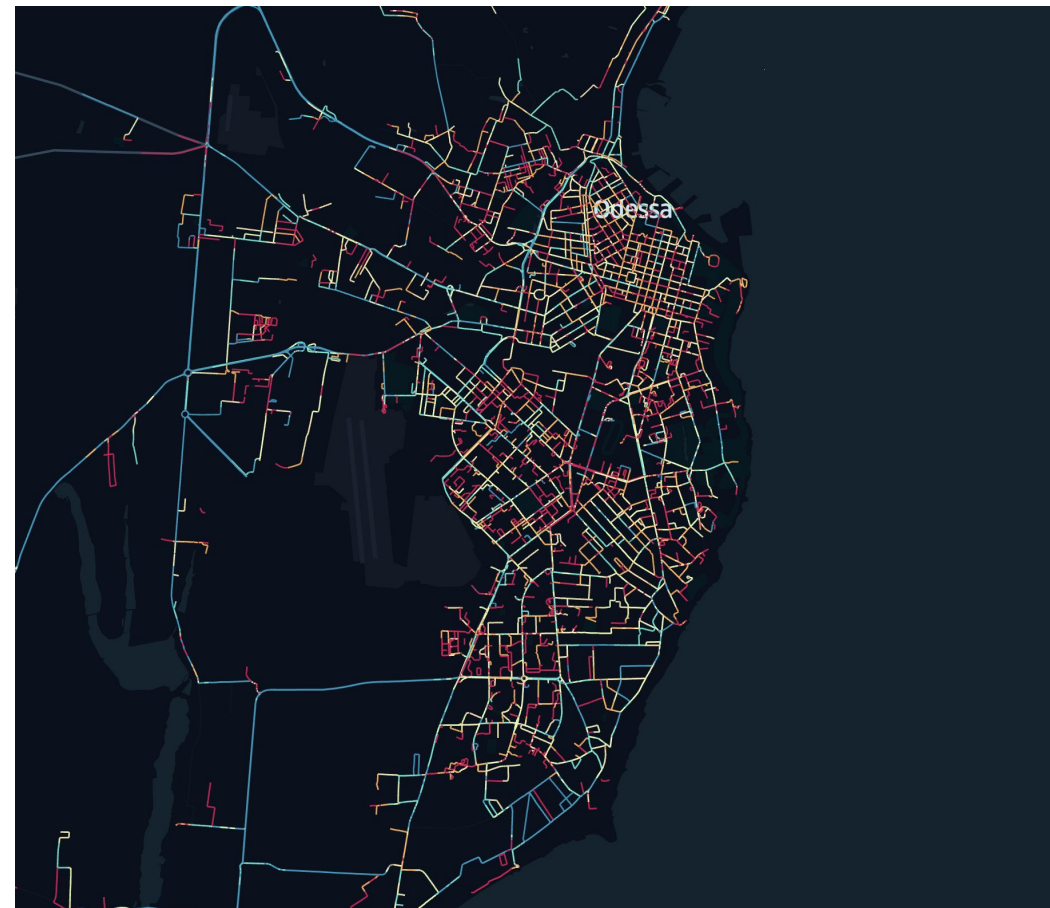
Id – унікальний ідентифікатор ордера (замовлення, створеного пасажиром).

running_time – дата та час коли почалась поїздка.

completed_time – дата та час коли завершилася поїздка.

route_distance_km – відстань в кілометрах від початку маршруту до місця призначення.

delta_time – фактичний час, за який завершилась поїздка.



Файл **nodes.csv**** містить дані:

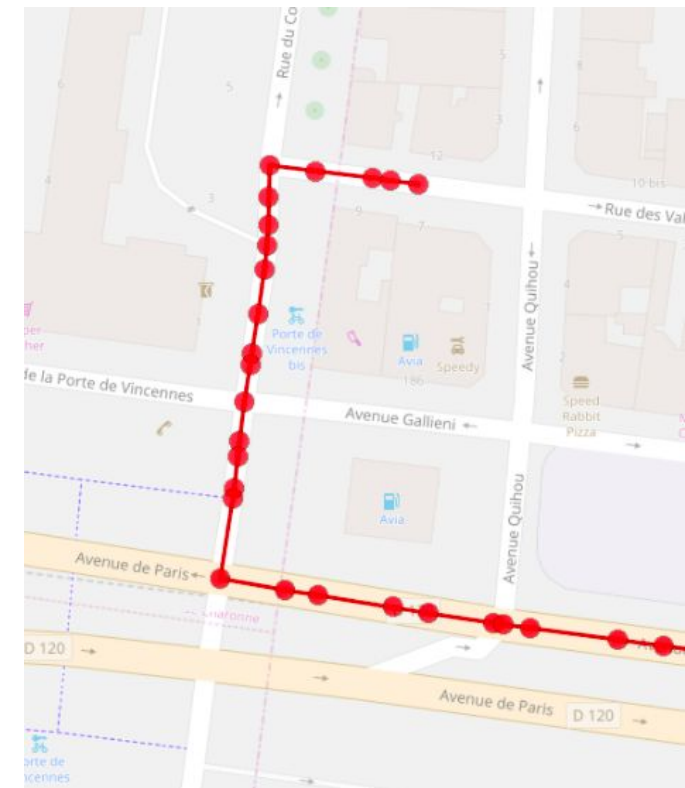
Id – унікальний ідентифікатор ордера (замовлення, створеного пасажиром).

node_start – початкова node.*

node_finish – кінцева node.*

distance – відстань (проміжок) між node-ми в метрах.

speed – опосередкована швидкість інших автомобілів на цьому проміжку до моменту створення замовлення.



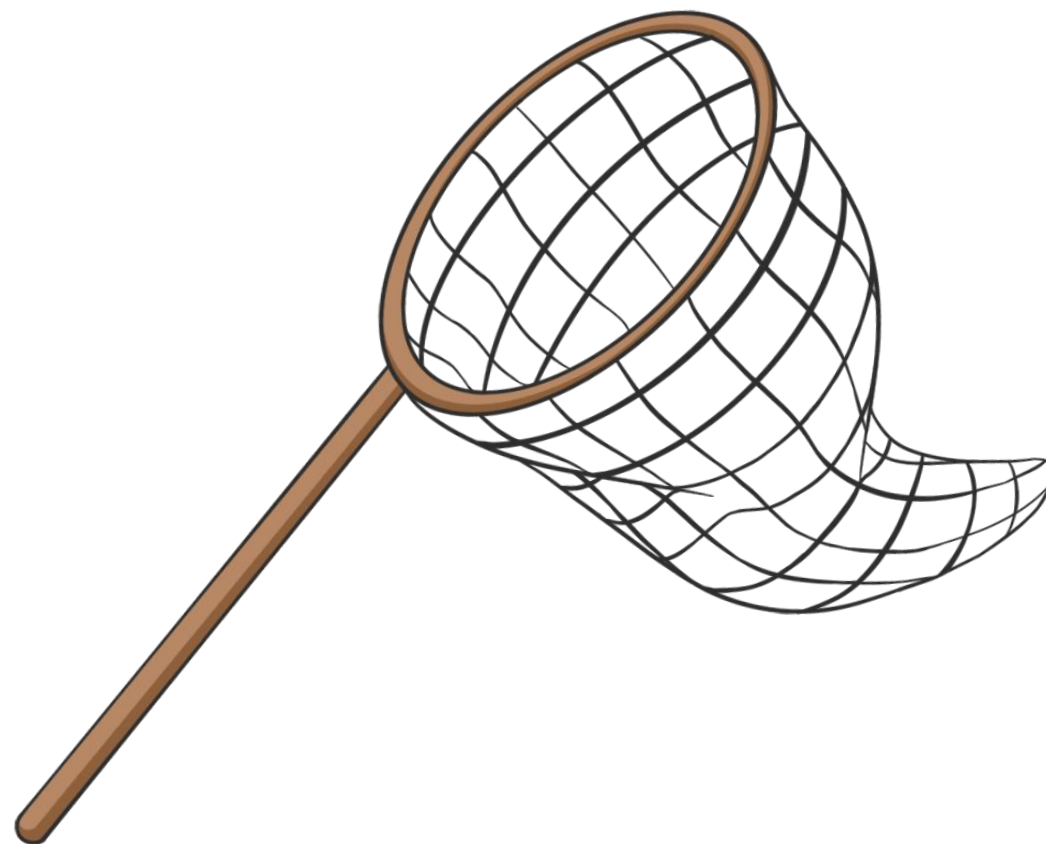
* *Node (вузол) – є одним з основних елементів моделі даних OpenStreetMap. Він складається з єдиної точки в просторі, визначеної її широтою, довготою та node id (ідентифікатором вузла)*

** *Для одного ордера з orders.csv буде існувати декілька записів в nodes.csv, які характеризують маршрут поїздки, розділений на менші прямі проміжки. Вони, своєю чергою, характеризуються розташуванням (початком та кінцем проміжку), дистанцією та опосередкованою швидкістю, за яку інші водії до моменту розміщення замовлення проїжджали цей проміжок.*

Використовуйте додаткові дані

Для вирішення завдання радимо використовувати відкриті джерела: [OpenStreetMaps](#), дані про погоду в той час в цьому місці та будь-які інші, що можуть впливати на транспортну ситуацію в місті та допоможуть підвищити точність прогнозування.

Пам'ятайте, що використовувати можна тільки ті дані, які можливо отримати до моменту створення замовлення (не можна заглянути у майбутнє).



Критерії оцінювання

60 балів – точність моделі

30 балів – застосовність моделі

10 балів – презентація



Критерії оцінювання

Точність моделі

Контрольна перевірка моделей буде проводитися на нових тестових даних, які будуть надані учасникам наприкінці хакатону. Для перевірки точності буде використовуватись RMSE метрика.

60
балів

Критерії оцінювання

Застосовність моделі

У критерії застосовності моделі входять:

- доступність відкритих джерел;
- швидкість прогнозування на тестових даних;
- відсутність “заглядань у майбутнє” та будь-які інші причини, через які модель не можна застосувати на практиці.

** Як приклад, використання в моделі відкритих даних, які ще не існують на момент створення замовлення пасажиром.*

30
балів

Критерії оцінювання

Презентація

Рекомендований час презентації 5-7 хвилин.

Презентація повинна містити:

- назву вашої команди та перелік учасників;
- перелік додаткових джерел, які ви використовували;
- перелік додаткових фіч з поясненням;
- найменування моделі, яку ви використали;
- ваші результати на тренувальній та тестовій вибірках;
- аналіз результатів моделі.

10
балів

**Бажаємо творчого
настрою та успіхів у
вирішенні завдання**

Let's drive technology!

