Проект – «Определение риска ДТП»

Нам нужно создать систему, которая могла бы оценить риск ДТП по выбранному маршруту движения. Под риском понимается вероятность ДТП с любым повреждением транспортного средства. Как только водитель забронировал автомобиль, сел за руль и выбрал маршрут, система должна оценить уровень риска. Если уровень риска высок, водитель увидит предупреждение и рекомендации по маршруту.

Наша задача:

1. Создать модель предсказания ДТП. Обязательное условие — учесть фактор возраста автомобиля.
2. На основе модели исследовать основные факторы ДТП. Понять, помогут ли результаты моделирования и анализ важности факторов ответить на вопросы: возможно ли создать адекватную системы оценки водительского риска при выдаче авто? Какие ещё факторы нужно учесть? Нужно ли оборудовать автомобиль какими-либо датчиками или камерой?

В ходе проекта мы:

1. Подключились к базе данных. Загрузили таблицы sql. (sqlalchemy)
2. Провели статистический анализ факторов ДТП. (pandas, pivot\_table, matplotlib)
3. Определили какие признаки будем использовать при обучении, создали sql запрос для итоговой таблицы.
4. Провели предобработку данных в итоговой таблице.
5. Подготовили выборки для обучения. (train\_test\_split, OneHotEncoder, StandardScaler)
6. Обучили несколько моделей и выбрали лучшую. (CatBoostClassifier, RandomForestClassifier, torch, GridSearchCV)
7. Проанализировали важность основных факторов, влияющих на вероятность ДТП
8. Сделали вывод и сформировали рекомендации.