

УДК: 004.942

**Название:** Прогнозирование посещаемости торговых точек на основе транзакционных данных.

**Авторы:**

Шиков Е.Н, Университет ИТМО, г. Санкт-Петербург;

**Научный руководитель:** Боченина К.О., Университет ИТМО, г. Санкт-Петербург

**Тезис доклада:**

Для коммерческих предприятий выбор местоположения играет решающую роль в успешном развитии и окупаемости. Быстрое развитие городов несет в себе большую неопределенность при выборе подходящего местоположения для нового заведения, поскольку следует учитывать множество сложных нелинейных факторов. В целом, существующие подходы основаны на данных, полученных вручную из опросов и отчетов, или косвенных данных, основанных на местоположении, которые не отражают реальные схемы расходов клиентов. Используя реальные транзакционные данные, можно привнести новый взгляд на эту проблему, поскольку эта информация напрямую отражает факты и объемы покупок в этой области.

Т.к. в исходных данных известно положение только 4% торговых точек, были разработаны методы для оценки координат точек из транзакционных данных. В качестве метрики использовалось среднее медианное расстояние между расчетным и реальным положениями в километрах. После инициализации было получено значение ошибки 0.41 км, после процедуры многомерной оптимизации координат точек - 0.25 км. Также была проанализирована зависимость погрешности от степени вершины и был сделан вывод, что по мере увеличения количества связей между точками ошибка определения местоположения уменьшается.

После успешного завершения процедуры определения координат, были сгенерированы признаки и обучена модель XGBoost. Коэффициент детерминации потока клиентов на тестовой выборке составил 0,68. Также был проведен дополнительный анализ важности характеристик и качества прогноза для каждой финансовой категории в отдельности.

Автор \_\_\_\_\_ / Шиков Е.Н. /

Научный руководитель \_\_\_\_\_ / Боченина К.О. /

Директор мегафакультета ТИнт \_\_\_\_\_ / Бухановский А.В. /