УДК 656.7.08

Успенская Александра Николаевна, студент,
ФГБОУ ВО «Санкт-Петербургский государственный университет гражданской авиации имени Главного маршала авиации А. А. Новикова»
Uspenskaya Alexandra Nikolaevna, Student,
St. Petersburg State University of Civil Aviation named after Chief Marshal of Aviation A. A. Novikov

Колчанова Дарья Дмитриевна, студент,
ФГБОУ ВО «Санкт-Петербургский государственный университет гражданской авиации имени Главного маршала авиации А. А. Новикова»
Kolchanova Daria Dmitrievna, Student,
St. Petersburg State University of Civil Aviation named after Chief Marshal of Aviation A. A. Novikov

Научный руководитель:
Соколов Олег Аркадьевич, к.т.н., доцент,
ФГБОУ ВО «Санкт-Петербургский государственный университет гражданской авиации имени Главного маршала авиации А. А. Новикова»
Sokolov Oleg Arkadevich,
Candidate of Technical Sciences(PhD in Engineering), Associate Professor,
St. Petersburg State University of Civil Aviation named after Chief Marshal of Aviation A. A. Novikov

ПРЕДИКТИВНАЯ АНАЛИТИКА ЗАДЕРЖЕК РЕЙСОВ: МЕТОДОЛОГИЯ ФОРМИРОВАНИЯ ПРИЗНАКОВОГО ПРОСТРАНСТВА И КРИТЕРИИ ОЦЕНКИ КАЧЕСТВА МОДЕЛИ
PREDICTIVE ANALYTICS OF FLIGHT DELAYS: METHODOLOGY OF FEATURE SPACE FORMATION AND MODEL QUALITY ASSESSMENT CRITERIA

Аннотация. В статье рассматривается методология построения прогнозных моделей задержек рейсов на основе эксплуатационных данных. Описана структура признакового пространства, включающая параметры расписания, метеорологические характеристики и данные о состоянии инфраструктуры. Проанализированы критерии оценки качества прогнозных моделей применительно к задачам операционного управления и обоснована необходимость различения статистической и операционной точности прогноза. Сформулированы практические требования к организации данных и процедуре эксплуатации модели.

Abstract. The article discusses the methodology for building predictive models of flight delays based on operational data. The structure of the feature space is described, including schedule parameters, meteorological characteristics, and infrastructure condition data. Quality criteria for predictive models applicable to operational management tasks are analyzed, and the need to distinguish between statistical and operational forecast accuracy is substantiated. Practical requirements for data organization and model operation procedures are formulated.

Ключевые слова: Предиктивная аналитика, задержка рейса, машинное обучение, операционное управление, качество прогноза, большие данные, гражданская авиация.

Keywords: Predictive analytics, flight delay, machine learning, operational management, forecast quality, big data, civil aviation.

Задержки рейсов остаются одним из основных факторов снижения эффективности деятельности авиаперевозчиков и аэропортовых операторов: они приводят к росту эксплуатационных расходов, нарушению стыковок и снижению удовлетворённости пассажиров. Организация воздушного движения и планирование расписания регламентируются правилами аэронавигационного обслуживания [1] и федеральными правилами использования воздушного пространства Российской Федерации [2], однако даже корректно составленное расписание не гарантирует соблюдения плановых времён операций. Традиционное планирование опирается на нормативные значения и средние исторические показатели, что не позволяет учитывать текущую операционную обстановку. Предиктивная аналитика предлагает иной подход: на основе накопленных эксплуатационных данных строится модель, оценивающая вероятность и величину задержки с заблаговременностью, достаточной для принятия управленческих решений.

Первым этапом построения модели является формирование признакового пространства. Целесообразно выделять четыре группы признаков. К первой относятся признаки расписания: плановое время вылета и прилёта, тип воздушного судна, принадлежность рейса к узловому или транзитному типу, интервал между последовательными рейсами одного борта. Ко второй – метеорологические признаки: видимость, нижний край облачности, интенсивность осадков, боковая составляющая ветра, прогноз на горизонт от одного до шести часов; метеофакторы действуют нелинейно, поэтому наряду с текущими значениями включают и скорость ухудшения условий. К третьей группе относятся признаки состояния инфраструктуры: загрузка взлётно-посадочной полосы и стоянок, доступность выходов, интенсивность наземного обслуживания. К четвёртой – каскадные признаки: задержка предшествующих рейсов того же борта, состояние обратного рейса, накопленная задержка в цепочке маршрута. Именно каскадные признаки обладают наибольшей информативностью на горизонте от двух до шести часов до вылета: наступившая в предыдущем пункте задержка с высокой вероятностью транслируется на последующие рейсы и известна заблаговременно.

Вторым этапом является выбор архитектуры модели и критериев её оценки. Применяются как классические алгоритмы – линейная регрессия, градиентный бустинг над деревьями решений [3], – так и рекуррентные нейронные сети; при хорошо сформированном признаковом пространстве бустинг нередко обеспечивает качество, сопоставимое с нейросетевыми моделями, при меньших затратах на сопровождение и большей интерпретируемости. При оценке качества необходимо различать статистическую точность, измеряемую средней абсолютной ошибкой прогноза, и операционную пригодность, определяемую полезностью прогноза для диспетчера: модель с меньшей средней ошибкой может уступать модели с большей, если последняя точнее предсказывает превышение порогового значения задержки, по которому принимаются решения. Различие между двумя классами критериев целесообразно фиксировать явно (табл. 1).

Поэтому наряду со статистическими метриками применяют операционные: долю корректно предсказанных критических ситуаций, частоту ложных тревог и средний выигрыш времени при превентивных мерах. Порог предупреждения следует выбирать исходя из цены ошибки обоих типов для конкретного аэропорта: его чрезмерное снижение увеличивает полноту, но быстро обесценивает прогноз в глазах персонала из-за частых ложных тревог.

Третьим этапом является организация данных и процедуры эксплуатации модели. Основные затруднения связаны не с алгоритмами, а с качеством данных: фиксация времён операций ведётся разными службами в разных системах, единицы измерения и часовые пояса не согласованы, часть событий регистрируется вручную. Обязательным требованием является единый справочник событий и регламент фиксации временных отметок; значительная часть

работ приходится на нормализацию исторических массивов. Не менее важно определение процедуры переобучения: распределение задержек меняется при изменении расписания, состава флота и сезонной нагрузки, поэтому модель должна периодически переобучаться на свежих данных с контролем деградации качества и автоматическим контролем статистик входных данных.

Отдельного рассмотрения заслуживает встраивание прогноза в рабочий процесс. Прогностическая оценка в диспетчерском интерфейсе меняет характер решений: возможны превентивные действия – заблаговременное изменение слота, перераспределение бригад, уведомление смежных служб [4]. Ответственность за решение сохраняется за человеком, а модель выступает рекомендательным инструментом; при этом прогноз должен сопровождаться объяснимыми факторами вклада, поскольку предупреждение без указания причин с трудом принимается персоналом. Ограничением применимости остаётся зависимость качества моделей от данных конкретного авиапредприятия, что определяет целесообразность поэтапного внедрения.

Таким образом, предиктивная аналитика задержек представляет собой последовательность организационно-технических решений, в которой алгоритмическая составляющая занимает подчинённое место по отношению к качеству данных и интеграции модели в процесс принятия решений [5].

Таблица 1

Критерии оценки качества прогнозной модели задержек рейсов

Группа критериев	**Показатель**	**Что характеризует**
Статистические	Средняя абсолютная ошибка	среднюю величину отклонения прогноза от фактической задержки
Статистические	Средняя абсолютная процентная ошибка	относительную точность прогноза для рейсов разной продолжительности
Операционные	Полнота по классу критических задержек	способность модели обнаруживать ситуации, требующие вмешательства
Операционные	Точность по классу критических задержек	обоснованность выдаваемых предупреждений
Операционные	Частота ложных тревог	дополнительную нагрузку на диспетчерский персонал
Операционные	Средний выигрыш времени	эффект превентивных мер при заблаговременном реагировании

Список литературы:

1. Международная организация гражданской авиации. Doc 4444. Правила аэронавигационного обслуживания. Организация воздушного движения. – Монреаль: ИКАО, 2016.

2. Федеральные правила использования воздушного пространства Российской Федерации (утв. Постановлением Правительства РФ от 11.03.2010 № 138), актуальная редакция.

3. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. – 2nd ed. – New York: Springer, 2009. – 745 p.

4. Предиктивная аналитика в управлении операционной деятельностью авиакомпаний: сборник научных трудов / под ред. Е. В. Соколовой. – М.: МАИ, 2023.

5. Современные методы анализа больших данных на воздушном транспорте: учебное пособие / под ред. Д. А. Крылова. – СПб.: СПбГУ ГА, 2022.