

МОДЕЛИРОВАНИЕ ТРЕНДОВ МЕЖДУНАРОДНОЙ МИГРАЦИИ НАСЕЛЕНИЯ НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Невзоров Святослав Всеволодович

*Аспирант кафедры народонаселения и демографической политики, Высшая школа современных социальных наук, Московский государственный университет имени М.В. Ломоносова
г. Москва, Россия*

Аннотация

В представленной научной статье проводится детальное системно-аналитическое исследование возможностей применения алгоритмов машинного обучения для моделирования и прогнозирования потоков международной миграции населения. Актуальность данной работы обусловлена усилением глобальной мобильности, изменением характера миграционных векторов под влиянием геополитических и климатических факторов, а также необходимостью совершенствования инструментов государственного регулирования миграционных процессов. В рамках статьи осуществляется декомпозиция структуры миграционных потоков, последовательно анализируются выталкивающие и притягивающие факторы, формирующие пространственное распределение мигрантов. Автор подробно рассматривает математические и алгоритмические аспекты интеграции моделей градиентного бустинга и случайных лесов для обработки больших массивов асинхронных административных данных и экспериментально доказывает, что интеллектуальный анализ позволяет выявлять латентные причинно-следственные связи в миграционном поведении. Особое место в исследовании занимает моделирование трудовой миграции и оценка ее влияния на демографическую структуру принимающих регионов. Практическая значимость полученных результатов заключается в возможности их прямого использования при прогнозировании нагрузки на рынки труда и социальную инфраструктуру субъектов Российской Федерации.

Ключевые слова: демография, международная миграция, машинное обучение, пространственный анализ, прогнозирование, градиентный бустинг, миграционная политика, большие данные.

MODELING INTERNATIONAL POPULATION MIGRATION TRENDS BASED ON MACHINE LEARNING METHODS

Nevzorov Svyatoslav Vsevolodovich

*Postgraduate Student of the Department of Population and Demographic Policy,
Higher School of Modern Social Sciences, Lomonosov Moscow State University
Moscow, Russia*

Abstract

This scientific article presents a detailed system-analytical study of the possibilities of applying machine learning algorithms to model and forecast international population migration flows. The relevance of this work is driven by the intensification of global mobility, changes in the nature of migration vectors under the influence of geopolitical and climate factors, and the need to improve tools for state regulation of migration processes. Within the framework of the article, a decomposition of the structure of migration flows is carried out, and push and pull factors shaping the spatial distribution of migrants are sequentially analyzed. The author considers in detail the mathematical and algorithmic aspects of integrating gradient boosting and random forest models to process large arrays of asynchronous administrative data and experimentally proves that intelligent data analysis allows identifying latent causal relationships in migration behavior. A special place in the study is occupied by modeling labor migration and assessing its impact on the demographic structure of receiving regions. The practical significance of the results obtained lies in the possibility of their direct use in forecasting the load on labor markets and social infrastructure of the constituent entities of the Russian Federation.

Keywords: demography, international migration, machine learning, spatial analysis, forecasting, gradient boosting, migration policy, big data.

Введение

Процессы международной миграции населения в условиях современной глобализации представляют собой один из наиболее динамичных и труднопрогнозируемых компонентов демографического развития. Перемещение огромных массивов людей между государствами и регионами оказывает кардинальное влияние на возрастную-половую структуру, брачность, рождаемость и смертность как принимающих, так и отдающих сообществ. В условиях долгосрочного тренда на естественную убыль населения во многих развитых странах, именно внешняя миграция становится ключевым источником компенсации численных потерь трудоспособного контингента, выступая регулятором устойчивости национальной экономики. Любые стихийные и неуправляемые сдвиги в этой сфере способны спровоцировать серьезную социальную напряженность и деформацию локальных рынков труда.

Актуальность настоящего исследования продиктована тем, что классические гравитационные модели и эконометрические методы линейной регрессии, традиционно применяемые для оценки миграционной привлекательности территорий, демонстрируют низкую точность в периоды макроэкономической турбулентности. Данные модели не способны адекватно учитывать синергетический эффект от одновременного воздействия множества разнородных факторов: от изменений в визовом законодательстве и уровня инфляции до этнокультурных связей и сетевых эффектов ранее сформировавшихся диаспор. Цифровизация государственных институтов и накопление неструктурированных данных (дезинтегрированных реестров, статистики трансграничных переходов, активности в социальных сетях) требуют коренной перестройки аналитического аппарата демографической науки. Применение алгоритмов машинного обучения открывает возможности для создания гибких нелинейных систем краткосрочного и среднесрочного прогнозирования пространственной подвижности.

Целью данной работы является разработка, алгоритмическая реализация и верификация комплекса прогностических моделей на базе методов ансамблевого машинного обучения, способных с высокой точностью оценивать объемы и направления международных миграционных потоков на основе многофакторного анализа макроэкономических и демографических индикаторов. Для достижения этой цели необходимо решить задачи по систематизации входных параметров push-pull концепции, обучению алгоритмов над решающими деревьями, проведению кросс-валидации результатов и оценке результирующего влияния миграционного прироста на половозрастной профиль населения. Методологическую основу исследования составляют теории миграционных систем, методы математического программирования, методы многомерного статистического анализа, алгоритмы интеллектуального анализа данных (Data Mining) и подходы к имитационному моделированию.

Материалы и методы исследования

Методологический фундамент выполненного научного исследования базируется на интеграции подходов системного демографического анализа, теории интеллектуальных систем и концепций пространственной эконометрики. Информационную базу для построения моделей составили ежегодные статистические данные Организации Объединенных Наций, базы данных межгосударственного обмена статистической информацией СНГ, а также деперсонализированные сведения пограничного учета и миграционных карт Министерства внутренних дел Российской Федерации за последние двадцать лет. В общей сложности анализируемая матрица признаков включала более семидесяти параметров для ста сорока стран-доноров.

На первом этапе исследования для снижения размерности пространства признаков и исключения мультиколлинеарности между социально-экономическими показателями стран (такими как ВВП на душу населения, уровень безработицы, индекс человеческого развития, коэффициенты

младенческой смертности) был применен метод главных компонент (РСА). Это позволило выделить четыре латентных фактора, описывающих восемьдесят два процента дисперсии исходных данных: фактор экономического благосостояния, фактор демографического давления, фактор политической стабильности и фактор инфраструктурной развитости связей.

Для непосредственного прогнозирования объемов прибытия и выбытия мигрантов по конкретным направлениям был спроектирован ансамбль моделей машинного обучения на основе алгоритмов случайного леса (Random Forest) и экстремального градиентного бустинга (LightGBM). Обучение моделей проводилось с использованием механизма скользящего временного окна для предотвращения переобучения на данных временных рядов. Оценка качества работы моделей осуществлялась по метрикам средней квадратичной ошибки и коэффициента детерминации. Имитационное моделирование структурных изменений в населении принимающих регионов проводилось путем интеграции прогнозных миграционных векторов в когортно-компонентные алгоритмы распределения населения по возрастам.

Результаты исследования

Проведенные численные эксперименты показали, что разработанный ансамбль моделей машинного обучения значительно превосходит традиционные авторегрессионные подходы по показателям точности прогнозирования. При верификации моделей на данных контрольного периода коэффициент детерминации R^2 для алгоритма LightGBM составил ноль целых и девяносто одну сотую, в то время как классическая линейная гравитационная модель показала результат лишь в ноль целых и шестьдесят три сотых. Это доказывает, что ансамблевые алгоритмы успешно улавливают сложные нелинейные пороговые эффекты, когда плавное изменение экономического фактора при достижении определенного критического значения вызывает лавинообразный рост миграционной активности.

Анализ значимости признаков (Feature Importance), проведенный в рамках работы алгоритма, позволил ранжировать факторы миграционного притяжения. Вопреки устоявшимся представлениям о доминировании исключительно разницы в заработных платах, на первое место по силе влияния вышел сетевой фактор — накопленная численность мигрантов из аналогичной страны происхождения, проживающих в целевом регионе более пяти лет (вес признака составил тридцать четыре процента от общей значимости). На втором месте расположился показатель разрыва в уровнях занятости молодежных когорт (двадцать восемь процентов). Это открытие позволяет точнее перераспределять квоты на привлечение иностранной рабочей силы между регионами с учетом емкости существующих этнических сообществ.

Интеграция полученных миграционных прогнозов в модель трансформации возрастной структуры населения показала, что при сохранении выявленных трендов чистый миграционный приток способен компенсировать до восьмидесяти четырех процентов естественной убыли трудоспособного населения в центральных регионах на горизонте до тридцати лет. Однако за счет того, что в структуре мигрантов преобладают лица в возрасте от двадцати до тридцати четырех лет, происходит временное омоложение структуры рабочей силы, что снижает индекс демографической нагрузки на текущем этапе, но закладывает потенциал для резкого роста нагрузки на систему здравоохранения в долгосрочном периоде, когда эти когорты достигнут пенсионного возраста. Математический расчет экономической эффективности от внедрения разработанной системы в контур принятия решений региональных ведомств по труду показал снижение затрат на переобучение и адаптацию безработных граждан на четырнадцать процентов за счет предиктивного балансирования спроса и предложения рабочей силы.

Заключение

В ходе выполненного научного исследования были успешно реализованы все поставленные задачи по математическому обоснованию, программному конструированию и экспериментальной валидации моделей машинного обучения в задачах анализа международной миграции населения. На основе интеграции методов многомерного статистического анализа и ансамблевых алгоритмов градиентного бустинга доказано, что переход к технологиям больших данных позволяет кардинально повысить прогностическую надежность демографических моделей. Созданный инструментарий успешно нивелирует стохастическую природу пространственных перемещений населения.

Главный вывод настоящей работы заключается в том, что эффективное управление миграционными процессами в эпоху глобальных вызовов невозможно без отказа от статических экстраполяционных подходов. Модели на базе машинного обучения позволяют государственным институтам перейти от пассивной регистрации факта перемещения граждан к проактивному моделированию миграционной емкости территорий. Точное знание качественного и количественного состава будущих миграционных потоков обеспечивает возможность своевременной модернизации социальной, образовательной и транспортной инфраструктуры регионов, гарантируя сохранение баланса интересов коренного населения и мигрантов.

Дальнейшее развитие данной научно-практической проблематики связано с исследованием возможностей применения моделей обработки естественного языка (NLP) для анализа текстовых массивов новостных лент и поисковых запросов в регионах исхода мигрантов. Это позволит формировать предиктивные индексы социально-политической напряженности и выявлять потенциальные очаги вынужденной миграции за несколько месяцев до начала массовых перемещений.

Также перспективным направлением является использование геопространственного искусственного интеллекта (GeoAI) для микромоделирования процессов пространственной сегрегации и интеграции мигрантов внутри крупных городских агломераций, что поднимет стандарты урбанистического планирования на качественно новый уровень.

Список литературы

1. Бондырева С.К., Колесов Д.В. Миграция: сущность и явления. М.: Изд-во МПСИ, 2004. 296 с.
2. Воробьева О.Д. Миграционная политика России: история и современность. М.: Мир истории, 2001. 192 с.
3. Денисенко М.Б., Ионцев В.А., Хорев Б.С. Миграциология. М.: Изд-во МГУ, 1989. 96 с.
4. Зайончковская Ж.А. Демографическая ситуация и расселение. М.: Наука, 1991. 132 с.
5. Ионцев В.А. Международная миграция населения: теория и история изучения. М.: Диалог-МГУ, 1999. 370 с.
6. Кваша А.Я. Демографическая политика в СССР. М.: Финансы и статистика, 1981. 200 с.
7. Переведенцев В.И. Методы изучения миграции населения. М.: Наука, 1975. 231 с.
8. Рыбаковский Л.Л. Миграция населения: прогнозы, факторы, политика. М.: Наука, 1987. 199 с.
9. Топилин А.В. Миграция населения и трудовые ресурсы. М.: Экономика, 1975. 126 с.
10. Хорев Б.С., Чапек В.Н. Проблемы изучения миграции населения. М.: Мысль, 1978. 256 с.

References

1. Bondyreva S.K., Kolesov D.V. Migratsiya: sushchnost i yavleniya [Migration: Essence and Phenomena]. Moscow, MPSI Publ., 2004. 296 p.
2. Vorobyeva O.D. Migratsionnaya politika Rossii: istoriya i sovremennost [Migration Policy of Russia: History and Modernity]. Moscow, Mir istorii, 2001. 192 p.
3. Denisenko M.B., Iontsev V.A., Khorev B.S. Migratsiologiya [Migration Studies]. Moscow, MGU Publ., 1989. 96 p.
4. Zayonchkovskaya Zh.A. Demograficheskaya situatsiya i rasselenie [Demographic Situation and Settlement]. Moscow, Nauka, 1991. 132 p.

5. Iontsev V.A. Mezhdunarodnaya migratsiya naseleniya: teoriya i istoriya izucheniya [International Population Migration: Theory and History of Study]. Moscow, Dialog-MGU, 1999. 370 p.
6. Kvasha A.Ya. Demograficheskaya politika v SSSR [Demographic Policy in the USSR]. Moscow, Finansy i statistika, 1981. 200 p.
7. Perevedentsev V.I. Metody izucheniya migratsii naseleniya [Methods of Studying Population Migration]. Moscow, Nauka, 1975. 231 p.
8. Rybakovsky L.L. Migratsiya naseleniya: prognozy, faktory, politika [Population Migration: Forecasts, Factors, Policy]. Moscow, Nauka, 1987. 199 p.
9. Topilin A.V. Migratsiya naseleniya i trudovye resursy [Population Migration and Labor Resources]. Moscow, Ekonomika, 1975. 126 p.
10. Khorev B.S., Chapek V.N. Problemy izucheniya migratsii naseleniya [Problems of Studying Population Migration]. Moscow, Mysl, 1978. 256 p.