**Прогнозирование международной миграции с использованием Google трендов**

Международная миграция оказывает влияние на различные сферы экономики и общества. Мигранты прибывают с разными профессиональными навыками, позволяющими вносить значимый вклад в обогащение человеческого капитала принимающих стран. Однако, статистические данные о внешней миграции публикуются с большой задержкой, которая может достигать нескольких лет. Данная проблема не позволяет исследователям осуществлять своевременный анализ миграционных потоков и их влияние на демографические факторы и уровень жизни населения, различия в доходах и другие макроэкономические показатели.

С распространением информационных технологий и возможностью пространственной привязки поисковых запросов  открываются новые возможности для статистического учета перемещений. Изучение поведения пользователей в сети Интернет позволяет опередить публикационный период статистических национальных данных (Hauzenberger et al., 2022), а также спрогнозировать действия населения в ближайшей перспективе (Varian, 2014).

Google Trends Index[[1]](#footnote-1) является одним из самых популярных инструментов для анализа информационного следа пользователей различных специализированных онлайн-ресурсов. Эмпирические данные указывают, что мигранты все чаще ищут актуальную информацию о переезде в сети Интернет в стране их происхождения до отъезда (Fantazzini et al. 2021, Böhme et al., 2020, Wanner, 2021). Совокупную интенсивность поисковых запросов можно использовать для обогащения косвенных показателей изменений количества миграционных намерений.

Цель данной работы – разработать метод прогнозирования международных миграционных потоков из России в Германию с использованием имеющейся официальной статистики стран, а также поисковых запросов в сети Интернет. Для анализа была взята статистика миграционной службы Германии и Росстата за период 2011-2020, был применен метод увеличения частотности статистических данных, который позволил из годовых показателей получить месячную частотность. Для оценки множества поисковых запросов, описывающих намерение мигрировать, были применены лингвистические методы машинного обучения. В результате была разработана линейная регрессионная модель, позволяющая предсказывать миграционные потоки, опираясь на данные множества поисковых запросов в Google Trends. Хотя оцененные параметры модели не могут быть использованы с другим временным промежутком, в работе описан принцип получения прогноза в будущем. Было проведено сравнение прогнозов моделей на основе сезонного тренда (SARIMA) и линейной регрессионной модели с переменными запросов Google Trends. Преимуществом модели, основанной на данных Google Trends, сравнительно с моделью SARIMA является возможность учитывать структурные сдвиги в динамике миграции из-за шоков различной природы (Covid-19, изменение социально-экономического положения, кризисы и др.). Это можно заметить на рис.1, где видно падение миграционного потока на период Covid-19 в начале 2020 года по модели, оцененной с использованием Google Trends.

Применение описанных выше моделей позволяет получить оперативные ежемесячные данные о внешней миграции раньше, чем они будут опубликованы в официальных статистических источниках (наукастинг миграционных потоков). Так данные Росстата могут иметь задержку более 2-х лет, что может затруднять оценку экономических показателей, основанных на демографических сведениях. Более того, в некоторых случаях описанный подход позволяет предсказать миграционные намерения еще до момента миграции.

Описанные методы можно использовать как при исследовании других пар стран, так и для оперативного прогнозирования иных статистических данных. При использовании предложенного подхода важно принимать во внимание особенности методологического учета данных о миграции в разных странах. Можно также проводить дополнительные исследования в области выбора множества поисковых запросов, выбирая множества, минимизирующие дисперсию при прогнозировании.

**Рисунок 1. Количество мигрантов из РФ в Германию по месяцам (фактические и оцененные значения)**

*Источник: Данные Google Trends Index, ОЭСР, Yandex Wordstat*

*Примечание: Множество поисковых запросов (поисковых слов) сформировано на основе поисковой истории Яндекса за последний год.*

**Литература**

Böhme, M. H., Gröger, A., & Stoehr, T. (2020). Searching for a better life: Predicting international migration with online search keywords. *Journal of Development Economics*, *142*, 102347.

 Fantazzini, D., Pushchelenko, J., Mironenkov, A., & Kurbatskii, A. (2021). Forecasting Internal Migration in Russia Using Google Trends: Evidence from Moscow and Saint Petersburg. *Forecasting*, *3*(4), 774–803. <https://doi.org/10.3390/forecast3040048>

Hauzenberger, N., Huber, F., & Klieber, K. (2022). Real-time inflation forecasting using non-linear dimension reduction techniques. *International Journal of Forecasting*. <https://doi.org/10.1016/j.ijforecast.2022.03.002>

Varian, H., & Choi, H. (2009). Predicting the Present with Google Trends. *Economic Record*, *88*. <https://doi.org/10.2139/ssrn.1659302>

Wanner, P. How well can we estimate immigration trends using Google data?. Qual Quant 55, 1181–1202 (2021). <https://doi.org/10.1007/s11135-020-01047-w>

1. Google Trends Index: <https://trends.google.com/trends/?geo=RU> [↑](#footnote-ref-1)