

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра математического моделирования и анализа данных

Аннотация к дипломной работе

**Мультистрановой анализ типологии пандемии COVID-19 с
использованием методов статистического и машинного обучения**

Корниевич Александра Константиновна

Научный руководитель: кандидат физико-математических наук, доцент
Малюгин В. И.

Минск, 2022

РЭФЕРАТ

Дыпломная праца, 46 старонак, 20 малюнкаў, 6 табліц, 11 крыніц.

Ключавыя словы: COVID-19, МУЛЬТЫКРАІНАВЫ АНАЛІЗ, СТАТЫСТЫЧНЫЯ ІНДЫКАТАРЫ COVID-19, ПАНЭЛЬНЫЯ ДАДЗЕНЫЯ, КЛАСТЭРНЫ АНАЛІЗ.

Аб'ект даследавання — эпідэміялагічны працэс COVID-19 у мультыкраінавым аспекце.

Мэта работы — ранжыраванне краін еўрапейскага рэгіёну і пабудова статыстычных краінавых рэйтынгаў і мультыкраінавага інтэгральнага рэйтыngu інтэнсіўнасці праходжання эпідэмічнага працэсу COVID-19, а таксама ацэнка ўплыву пандэміі на эканамічныя паказчыкі краін.

Метады даследавання — метады машыннага навучання, метады кластарнага аналізу панэльных даных, метады пабудовы статыстычных рэйтынгаў.

Вынік — сфарміравана абнаўляемая база штодзённых статыстычных даных; сфарміравана прастора класіфікацыйных прыкмет і распрацаваны алгарытм кластарнага аналізу штомесячных і штодзённых панэльных даных; распрацаваны статыстычныя краінавыя рэйтынгі і інтэгральны мультыкраінавы інтэгральны рэйтынг інтэнсіўнасці праходжання эпідэмічнага працэсу COVID-19; ацэнены ўплыў пандэміі на тэмпы росту ВУП і ўзровень занятасці ў разгляданых краінах.

Вобласць прымянення — прымяненне метадаў аналізу панэльных дадзеных з кластарнай структурай для аналізу пандэміі COVID-19 у мультыкраінавым аспекце.

ABSTRACT

Graduate work, 46 pages, 29 drawings, 6 tables, 11 sources.

Keywords: COVID-19, MULTI-COUNTRY ANALYSIS, COVID-19 STATISTICAL INDICATORS, PANEL DATA, CLUSTER ANALYSIS.

Object of research — COVID-19 epidemiological process in a multi-country perspective.

Purpose of the work — ranking of countries in the European region and construction of statistical country ratings and a multi-country COVID-19 integrated ranking of the intensity of the epidemic, and assessment of the impact of the pandemic on the economic performance of countries.

Research methods — machine learning methods, methods for cluster analysis of panel data, methods for constructing statistical rankings.

Result — An updated database of daily statistics was generated; a space of classification attributes was created and an algorithm for cluster analysis of monthly and daily panel data was developed; statistical country ratings and an integral multi-country integral rating of the intensity of the epidemic COVID-19 were developed; the impact of the pandemic on GDP growth and employment rates in the countries in question was assessed

Scope — application of panel data analysis techniques with cluster structure to analyse the COVID-19 pandemic in a multi-country perspective.

РЕФЕРАТ

Дипломная работа, 46 страниц, 20 рисунков, 6 таблиц, 11 источников.

Ключевые слова: COVID-19, МУЛЬТИСТРАНОВОЙ АНАЛИЗ, СТАТИСТИЧЕСКИЕ ИНДИКАТОРЫ COVID-19, ПАНЕЛЬНЫЕ ДАННЫЕ, КЛАСТЕРНЫЙ АНАЛИЗ.

Объект исследования — эпидемиологический процесс COVID-19 в мультистрановом аспекте.

Цель работы — ранжирование стран европейского региона и построение статистических страновых рейтингов, мультистранового интегрального рейтинга интенсивности протекания эпидемического процесса COVID-19, а также оценка влияния пандемии на экономические показатели стран.

Методы исследования — методы машинного обучения, методы кластерного анализа панельных данных, методы построения статистических рейтингов.

Результаты — сформирована обновляемая база ежедневных статистических данных; сформировано пространство классификационных признаков и разработан алгоритм кластерного анализа ежемесячных и ежедневных панельных данных; разработаны статистические страновые рейтинги и интегральный мультистрановой интегральный рейтинг интенсивности протекания эпидемического процесса COVID-19; оценено влияние пандемии на темпы роста ВВП и уровень занятости в рассматриваемых странах.

Область применения — мультистрановой анализ пандемии COVID-19 на основе регулярно обновляемых панельных данных с кластерной структурой.

ВВЕДЕНИЕ

В данный момент, мир переживает пандемию нового и до этого неизвестного вируса COVID–19. Мир знает очень мало о данном вирусе, и информация появляется каждый день. Актуальная информация, мгновенные обновления, умение прогнозирования эпидемиологического процесса, анализ типологии протекания являются очень важными как для борьбы с вирусом, так и для мер предосторожности. В случае эпидемии нужно знать, сколько коек в больницах следует подготовить, сколько аппаратов ИВЛ произвести, на сколько месяцев сделать запасы средств индивидуальной защиты для медиков, какие меры эффективны в борьбе с эпидемией, а от каких следует отказаться. Должностные лица должны понимать, достаточно ли принятых мер, эффективны ли они, или следует придумать новые более эффективные запреты и ограничения. В идеале, модель должна настолько хорошо отражать реальность, чтобы по ней можно было увидеть силу действия каждой вновь принятой меры, и тогда можно было бы усиливать полезные меры и отменять решения, оказавшиеся бесполезными. Построение моделей эпидемиологического процессам должно основываться на всестороннем статистическом анализе важнейших его показателей с учетом характера его протекания и экономических последствий. В условиях пандемии, охвативших практически все страны мира большое значение имеет международном аспект данной проблемы.

В связи с актуальностью проблемы на просторах интернета появляется множество статей, связанных с влиянием COVID-19 на различные сферы жизни человечества. Например, есть статьи, в которых проводится эконометрический анализ случаев заболевания COVID-19, смертей и метеорологических факторов в Южной Азии [1]. Это первое исследование, в котором предпринята попытка изучить связь между подтвержденными случаями COVID-19, смертями, метеорологическими факторами и загрязнителем воздуха в некоторых странах Азии с использованием передовых эконометрических методов. Результаты исследования подтверждают, что существует сильная кросс-секционная зависимость и значительная корреляция между случаями COVID-19, смертями, метеорологическими факторами и загрязнителями воздуха для всех шести исследуемых стран Южной Азии [1]. Так же был проведен предварительный многострановой сравнительный анализ воздействия COVID-19 и стихийных бедствий в Индии, Японии, Филиппинах и США, который выявил необходимость улучшения общего состояния эвакуационных центров, чтобы побудить людей принять необходимые меры эвакуации без промедления и

беспокойства заразиться COVID-19. Это исследование было проведено для того, чтобы понять вопросы, проблемы и необходимые шаги по подготовке к будущим комплексным опасностям, которые нельзя исключать во время пандемий [2].

Целью данной работы является статистический анализ пандемии COVID- 19 в международном аспекте и оценка ее влияния на экономические показатели различных стран.

Для достижения данной цели решаются следующие задачи:

- разработка постоянно обновляемой базы данных для моделирования и анализа эпидемиологического процесса COVID-19 в международном аспекте;
- исследование типологии протекания эпидемиологического процесса COVID-19 в странах европейского региона и некоторых других стран с особенностями протекания эпидемии с помощью методов кластерного анализа панельных данных;
- оценка влияния эпидемиологического процесса COVID-19 на экономические показатели стран;
- разработка программного приложения, реализующего постоянно обновляемую базу данных, а также методику применения методов кластерного анализа для классификации панельных данных