Formulier Informatieverzoek Coronadata

Project Coronadata, datum 13 october 20

Nummer

INFORMATIEVERZOEK - VRAAG GESTUURDE

TERUG NAAR STATUSOVERZICHT

ONDERDEEL:
REFERENTIENUMMER:
ONTVANGSTDATUM:
NAAM:
ORGANISATIE:
FUNCTIE/ROL:
E-MAILADRES:
TELEFOONNUMMER:
HOOFDVRAAG:
DEELVRAAG 1:
DEELVRAAG 2:

CONTEXT:
RISICOANALYSE:

ADVIES WERKGROEPEN
PRIORITEIT:
STATUS:

Time Travel

ATIEVERZOEK - VRAAG-GESTUURDE CORONADATA

SOVERZICHT

BESCHRIJVING:
20211013-001
13th of october
5.1.2e
Delft University of Technology
5.1.2e
5.1.2e @tudelft.nl
06 5.1.2e
How can Machine-learning approaches help identify socio-spatial patterns in corona-related datasets (spread of corona virus, spread c impact of vaccination) over time by avoiding (pre-)selection? Which aproaches are suitable and what kind of insights can they produc on socio-spatial patterns generated by the projet help the medical community predict the spread of corona, COVID-19 and vaccinatior take timely and space-specific measures? For example: how does the fact of having an elevator (high-rise dense housing) influence th what measures can be taken to prevent/slow down the spread of the virus through temporary measures or architectural interventions
What socio-spatial disparities and patterns can we identify on neighborhood/urban/regional scales when using the Machine-learni during hte project on both social-spatial and to corona related datasets (spread of corona virus, spread of Covid-19 disease, impact cases of Germany and The Netherlands?
How can we derive a transferable and scalable framework from the analysis of the socio-spatial datasets of corona-related d Netherlands for use in other countries or at different scales (note: a similair research did start in Germany by the HafenCity U operation with the Robert Koch Institute)?

Background:
-Existing research on corona-related data tends to focus on studies on individual health risk factors and their identification. (For othe
as pneumonia and influenza the influence of spatial and socio-economic factors were analyzed in addition to individual risk.)
-Existing research focuses on predetermined data (epidemiological and socio-economic).
Aim:
-This study focuses on analyzing the spread of the corona virus, the COVID-19 disease and the impact of vaccination at the scale of n
the region.
-It uses Machine learning to display diverse non-pre-determined spatial, socio-cultural as well as COVID-19 related datasets in spac
overcome knowledge-related biasses.
-It therewith aims to uncover less known or expected relationships between the socio-spatial and corona-related data. These fin
instance take prevention measures, to set up in-depth research on these findings etc.
Methodology:
-Using Machine learning to cross open access socio-spatial data (some hundred indicators) with corona-related data to ider
demographic and cultural factors exploring how COVID-19 spreads, its development over time and what the implication for vaccinatio
Existing research on corona-related data tends to focus on studies on individual health risk factors and their identification. Thi
approach. It aims to investigate the relationship of a large collection of socio-spatial factors and different COVID-19 factors, like
hospitalization, or vaccination rates. In contrast to traditional correlation-based approaches, we aim at predicting such COVID-19 hi
already passed ("time travel") using supervised machine learning. A machine learning model trained for such an historic prediction ta
relationships between the input data (social-spatial factors) and the labels (COVID-19 factors) in a novel way. With the use of expla
can uncover such relationships for subsequent analysis.
-Using non-predetermined, multi-source and fine scale data sets to help reveal unexpected correlations between spatial forms, soc
spread of COVID-19 disease, the takeup of vaccination and their spread over time (complementing the results from the traditional me
-Analyzing computer generated data to assess the relevance and meaning of the findings
Goals:
-Draw conclusions from large multisource datasets at diverse scales and administrative levels to understand health-related factor
inequalities (e.g. density Inhabitants, commuting distance).
-Develop a transferrable and multi-scalar framework/methodology in which spatial factors (e.g. living close to a harbor, living in a sky
included in corona-related data analysis
-To support planners, decision-makers as well as civil society with evidence-based planning strategies to promote adaptive and liv
Results will be available within 6 months for assessment by the medical community in an interpretation session hostec
inviting relevant stakeholders.

1) The research project relies on the availability of socio-spatial and temporal data at the scale of the neighborhood that can be crosse
datasets: Most socio-spatial data is available open access. We also have access to corona-related data as long-term aggregates. The re
CBS microdata to be analyzed which was discussed with CBS
2) Patterns and relatiotionships uncovered in a data-driven fashion using machine learning can potentially be not meaningful. Before a
drawn from such results, they need to be externally verified with stakeholders.
3) As with most data-driven aproaches, there is a risk to uncover also meaningless correlations. We will take this into account when an
will verify and validate any results discovered from Machine Learning based analysis.
4) The scale and granularity of data available to us might simply not enough to uncover meaningful correlation between corona incide
factors.
5) There might be many factors muddying the observable relationships between socio-cultural data and corona incidents: personal mc
communiting, the progressing vaccination campaign (which we try to include but which is still hard to model), etc. We would like to
temporally fine-grained Covid-19 related datasets for Netherlands and Zuid Holland, including specifically Information on the (evolving
cases, hospitalization, intense care, death, vaccination rate and further Covid-19 related datasets as for example preliminary datasets
can help to improve the precision and quality of our analysis. To best address our research questions, temporal (daily, weekly or mont
scale levels are needed for the whole of the Netherlands and for South Holland in particular.

Communicatie werkgroep:
In later stadium

Vraagarticulatie werkgroep: Positief want
o Relevant en mooi onderzoek.
o Goed dat TU Delft een voorstel indient met mogelijkheden tot internationale vergelijking
o Biedt mogelijkheden tot meer dan statische kaartjes: Bijv geografische patronen en daarmee levert het gegarandeerd nieuwe i name een interessante onderzoeksvraag voor een volgende pandemie
o Suggestie om ook rioolwateronderzoek meenemen.
o Hebben de onderzoekers beschikking over contactnetwerken?

o Het heeft niet de hoogste urgentie heeft.

2

Intake informatieverzoek