インターネット利用の格差の研究

ーPythonを用いた回帰分析ー

#32

大泉高校Techcamp　佐々木俊輔

①要旨

新型コロナウイルスが地球規模で猖獗を極める現在、日本の教育現場ではオンライン授業が導入され、インターネット利用に関する格差が問題となっている。世界的にもアフリカや南アジアでインターネット利用率の低さが目立ち、先進諸国との格差は顕著である。そこで、国際間のインターネット利用の格差是正について研究した。この格差を考えるうえで注目すべき事柄を探るため、52本の論文のテキストマイニング結果等を参考に、8項目のデータを取得した。2変数の積項も追加して説明変数を44項目とし、Pythonを用いて線形系・決定木系の回帰分析をそれぞれ交差検証法とグリッドサーチを利用して行い、インターネット利用率を予測するモデルを構築した。SHAPモデルにより、所得や識字率、経済・教育分野の男女格差がインターネット利用率に影響すること、特に所得が与える影響が大きいことが分かったが、今後は所得格差の是正についても考究する必要がある。また説明変数を見直し、具体的な解決方法を示したい。

②研究背景

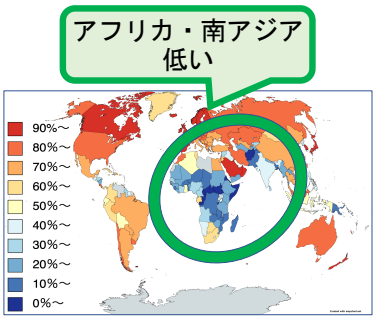

図1.インターネットを利用する個人の割合
ITU（2021）「World Telecommunication/ICT Indicators Database」をもとにMapChartを利用して作成

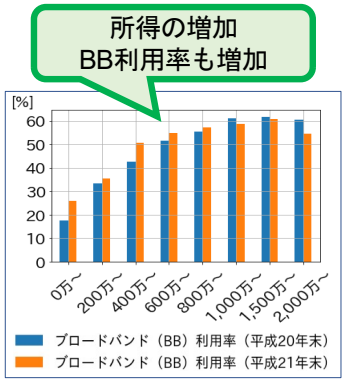

図2. 国内における所属世帯年収別
ブロードバンド利用率
総務省（2009）「平成21年通信利用動向調査」をもとに作成

③研究方法

取得　テキストマイニング『*Nature*』

加工　2変数の積項の追加
標準化

分析　回帰分析
SHAPモデル

income
education
white
women
household
student
infrastructure
nations
black
language
rural
gender
health
libraries
God
urban
library
literacy
site
men
reported
due
further
average
order
location
increasing
effort
addition
control
market
good
concept
people
place
resources
might
participants
least
form
price
training
project
life
very
better
work
after
during
findings
system
school
educational
school
system
connection
patterns
racial
process
up
set
standard
technology
technologies
organization
security
author
influence
policies
policy
interest
continue
take
time
since
since
since