

# Predicting & Understanding The Behavioural Footprint Of Older Adults In The UK

Thabsheera

30/05/2022

## Introduction

Technology had succeeded in increasing the life-span of human over the years. But not everyone is living a wealthy, healthy and a happy older life. So it's crucial for science to advance in order to improve the quality of life and not just adding years to life.

Ageing Intelligence is a new term coined by Nic Palmarini (Director of UK National Innovation Centre for Ageing - NICA). According to him ageing intelligence is a comprehensive data-driven approach, leveraging a combination of human knowledge and experience, data analysis, and real-world applied testing and research.

Analysing the trend and generating and amplifying insights from behavioural and market data is one of the key aspects of ageing intelligence. Behavioural data is a very important data irrespective of the domain. If it is an e-commerce market, behavioural data of customers can help in better recommendations and personalisations. When it comes to the behavioural data of older population in the country, it leads to numerous opportunity benefiting both the citizen and business. By knowing what an individual will be doing at a particular time of the day,

what is the future

how predicting a behaviour can improve lifestyle

## Project Aim

UK time survey data for 2014-15 has diary data for 4,000 households in the UK. Each diary data contains what that individual is doing at a particular time of the day. The high-level aim of this project is to predict what that person will be doing at a given time of the day. This project will require incorporating a classification model (like decision tree or random forest) into a time series component. Because diary data has time series aspect to it and the activities are categorical variables.

Along with predicting the activities of older adults at a given hour of the day, combining two machine learning techniques adds a high value to this project.

## Progress

The available data is obtained from UK time-survey data for the period of 2014-15. The data contains diary activities by around 8000 people from 4000 households in the UK. The data is already with a given weight for household and individual so that this will be comparable to the entire UK population. The granularity of the data is household, person, weekday/weekend, 10 minute interval of the day.

Basic EDA is conducted on the entire data to understand the distribution. Following are the major findings from the data:

1. Age group level
2. these kind of activities happen more during weekdays

3. People over 60 does more of this than other
4. Household without children tend to do these activities compared to adults only households.
5. Most folks enjoy doing self care but a few of this age group enjoy doing travelling more

To visualise the above findings, a R Shiny dashboard is built.

When it comes to behaviour classification, the data needs to be prepared in a certain way. First of all the important field like activity needed tweaking. The activities are micro behaviour like vehicle maintenance, dwelling repair, visiting historic site etc. There are 276 such behaviours which will increase the complexity of a classification model by will requiring 276 output groups. So these micro behaviours are then grouped into 48 groups. For example, activities/ behaviours like feeding the child, teaching the child etc is now under a group called 'child care'. This will reduce the complexity of the model.

The activities are recorded for every 10 minute interval of the day, this interval is then converted to hour of the day. So that instead of 144 time interval, 24 time intervals can be used for predicting the activities in that interval. The 10 minute interval is avoided since most of the activities overlap the intervals. To convert the data into hour level, the activity is replaced with the most frequent activity in that hour.

## **Work Plan**

Below is the rough work plan for the project

# Research Data Management Proforma

<br>

0. Project title, author, version and date		
Project:		
Author:	Version:	Date:
1. Description of the data		
<p><b>1.1 Type of study</b></p> <p><i>Up to three lines of text that summarise the type of study (or studies) for which the data are being collected.</i></p> <p><b>1.2 Assessment of existing data</b></p> <p><i>An explanation of the existing data sources that will be used by the research project, with references. Or an analysis of the gaps identified between the currently available and required data for the research.</i></p> <p><b>1.2 Types of data</b></p> <p><i>Types of research data to be managed in the following terms: quantitative, qualitative; generated from surveys, interviews, administrative records, simulations, images. Include the use of any code or software if it aids the creation or processing of the data.</i></p> <p><b>1.3 Format and scale of the data</b></p> <p><i>File formats, software used, number of records, databases, (in terms that are meaningful in your field of research). Do formats and software enable sharing and long-term validity of data? How large is the data going to be?</i></p>		
2. Data collection / generation		
<p><i>Focus on the good practice and standards for ensuring new data are of high quality and processing is well documented.</i></p> <p><b>2.1 Methodologies for data collection / generation</b></p> <p><i>How the data will be collected/generated and which community data standards (if any) will be used at this stage.</i></p> <p><b>2.2 Data quality and standards</b></p> <p><i>How consistency and quality of data collection / generation will be controlled and documented, through processes of calibration, repeat samples or measurements, standardised data capture or recording, data entry validation, peer review of data or representation with controlled vocabularies.</i></p>		
3. Data management, documentation and curation		
<p><i>Focus on principles, systems and major standards. Focus on the main kind(s) of study data. Give brief examples and avoid long lists.</i></p> <p><b>3.1 Managing, storing and curating data.</b></p> <p><i>Briefly describe how data will be stored, backed-up, managed and curated in the short to medium term (<a href="https://www.ncl.ac.uk/library/academics-and-researchers/research/rdm/working/">https://www.ncl.ac.uk/library/academics-and-researchers/research/rdm/working/</a>)</i></p> <p><b>3.2 Metadata standards and data documentation</b></p> <p><i>What metadata is produced about the data generated from the research? For example descriptions of data that enable research data to be used by others outside of your own team. This may include documenting the methods used to generate the data, analytical and procedural information, capturing instrument metadata alongside data, documenting provenance of data and their coding, detailed descriptions for variables, records, etc (see: <a href="http://www.dcc.ac.uk/resources/metadata-standards">www.dcc.ac.uk/resources/metadata-standards</a> / <a href="http://www.data-archive.ac.uk/create-manage/document">http://www.data-archive.ac.uk/create-manage/document</a>)</i></p>		
4. Data security and confidentiality of potentially disclosive information		

<br>