Early Diagnosis of Parkinson's Disease via keystrokes

Chin Min Tee
February 11, 2018

Readme

Environment:

Anaconda Navigator 1.6.12
Python 3.6
Jupyter notebook 5.2.1
Spyder 3.2.6

Special package needed:

xgboost

To prepare to produce csv files for jupyter notebook:

- Raw data cen be found:
  Tappy data:
  https://physionet.org/physiobank/database/tappy/Archived-Data.zip
  https://physionet.org/physiobank/database/tappy/Archived-users.zip
  NQdata:
  https://physionet.org/physiobank/database/nqmitcsxpd/neuroQWERTY.zip
- Raw data need to be in following directory:
  data/ArchivedUsers/User_xxx.txt (xxx is Tappy data user)
  data/TappyData/xxx_nnn.txt (nnn is file id for xxx user)
  data/MIT-CS1PD/GT_DataPD_MIT-CS1PD.csv
  data/MIT-CS1PD/data_MIT-CS1PD/*.csv (* is file identified in GT_DataPD_MIT-CS1PD.csv
  data/MIT-CS2PD/ GT_DataPD_MIT-CS2PD.csv
  data/MIT-CS2PD/data_MIT-CS2PD/*.csv (* is file identified in GT_DataPD_MIT-CS2PD.csv
  dfData/df.csv (file prododuced)
  dfData/nqdf.csv (file produced)
  dfData/*.csv (various files for exploratory)
- In same level as data directory, there should be:
  prepare.py (execute this to obtain df.csv
  prepareNQ.py (execute this to obtain nqdf.csv)
  constants.py
  funcs.py

mytappy.py

Jupyter notebook:

- ParkinsonDiseaseKeystroke.ipynb
- PD_exploration.ipynb

Report and proposal:

- ChinCapstoneProjReadme.docx (this document)
- ChinCapstoneProjReport.docx (project report)
- ChinCapstoneProposal.docx (project proposal)

Github:

- https://github.com/cteeeri/parkinson_keystroke.git