Current classification options

- PTB-XL (Original)
  - 1 source, the original PTB-XL
- 21,837 registers
- NORM, MI, STTC, CD, HYP
  - Criteria in next slide
- 74.56 % *

- PhysioNet/Computing in Cardiology Challenge 2020
  - 6 sources, including the 21,837 now reclassified PTB-XL registers
- 43,101 registers
- Urgency Level (1, 2, 3, 4)
  - As determined internally
- 49.37 % *

SNOMED CT Codes (111)

SNOMED International

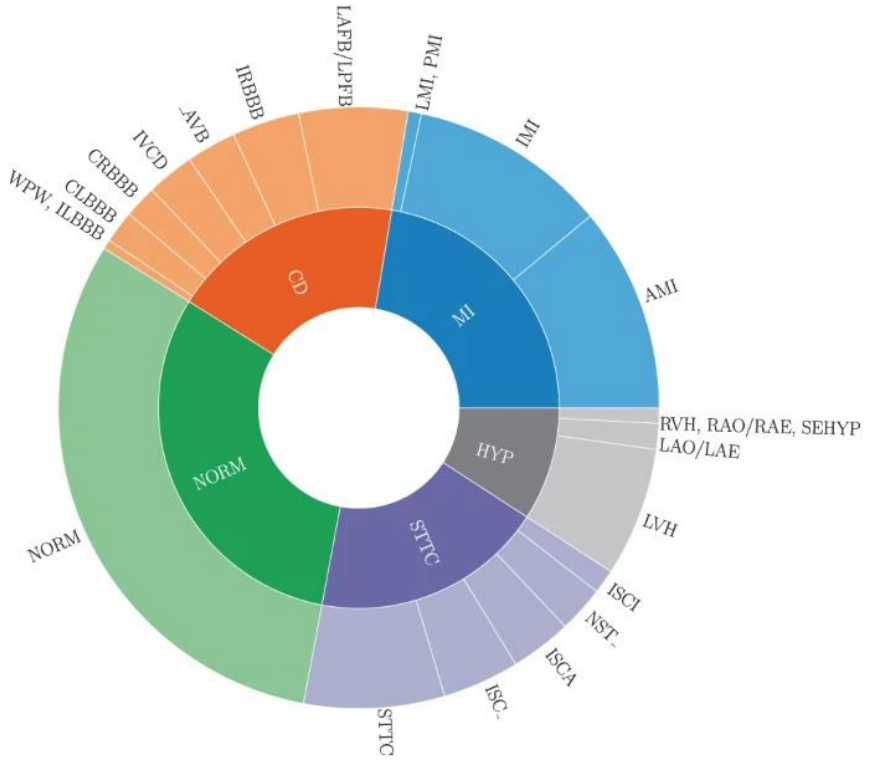
20.69 % * for 25 codes

* Validation set accuracy; test accuracy can be slightly lower. This is not for the final models, there are several ideas still untested.
On another note, the models for the PhysioNet datasets have only been tested with the 21,837 reclassified PTB-XL set.

PTB-XL Labelling

#Records	Superclass	Description
9528	NORM	Normal ECG
5486	MI	Myocardial Infarction
5250	STTC	ST/T Change
4907	CD	Conduction Disturbance
2655	HYP	Hypertrophy

Number of records for each of the labels.



Proportion of labels along their sublabels.

More detail in the following file:

(in Github : data/scp_statements.csv)

Microsoft Excel
ma Separated Valu

Some comments

- Many things still to test for
  - Regularizers
    - Dropout
    - L1
    - Where to put them and how they affect the results
  - Transfer Learning
  - Top n teams' architectures
  - Length of signal
  - LR adaptation
  - Weighted-by-class loss function
  - Metric to optimize for (maybe accuracy isn't the best but speficity or recall are)

- Using PTB-XL (Original) labels on the 43,101 records from PhysioNet
  - If there was a mapping from the PTB-XL original labels to each SNOMED CT Code, we could use any approach with the whole 43,101 registers.