Peer review

1. Nithya Mylakumar M, Sai Divya Teja Konda, Yasasree singam

5/2

* Goal
  + Malaria image classification
* Workflow
  + Data: from Kaggle; parasitized and uninfected
  + EDA: balanced
  + Data preprocess
    - \
  + Model
    - CNN: baseline model
    - Vgg10(CNN with 19 layers deep)
    - Resnet 50(residual neural network)
* What was done well?
  + \
* What do you learn from it?
  + New model algorithms

1. Chaeyeon Yim, Anna Jeffries, Angelica Gacis

5/2

* Goal
  + Modeling the likelihood of death within the first 24 houes
  + The role of bias according to ethnicity and gender
  + Critique current methods used by hospitals
* Workflow
  + Data: WiDS 2020 dataset(kaggle)
  + EDA
  + Data preprocess
    - remove categorical and ‘apache’ feature
    - encode ‘gender’ and ‘ethnicity’
    - for numeric data
      * impute NaN using SimpleImputer and mode
      * balanced using imbalance-learn
      * fill all NaN with 0
  + model
    - SVM
    - XGBoost(SMOTE to rebalance data)
    - Neural network(weighted vs unweighted)
* What was done well?
  + Beautiful plot during the EDA part
  + Balance imbalanced data
  + Hyperparameter tuning
  + Comparing the results of model with different dealing methods for numeric data
* What do you learn from it?
  + Preprocess is really important.
  + We could compare multiple preprocesses with their different results.