ETC3250 - Project

Di Cook and Souhaib Ben Taieb

Can you tell who is speaking, Di or Souhaib?

We have recorded Di and Souhaib saying different words: "regression", "date", "data", "explore", "learn", "model", "classify", "business", "bagging", "time", "analysis", "spline", "analytics" and "cluster". Each word was repeated multiple times. The full data set contains 122 audio segments (61 for both Di and Souhaib). The audio files were loaded into R using the `tuneR` package, and segments were created by down sampling the time series to provide an equal number of amplitude values for each word spoken. Amplitudes were also scaled to a mean 0 and sd 1.

Your task is to build a classifier for who is speaking. We have given you a training set of 80 audio segments which are fully labelled by word and speaker. Build your best classifier on this training data. Another 42 audio segments have been kept for testing your classifier. In this data set, the word and the speaker have been removed. You need to predict this set, and upload your predictions to kaggle to get an estimate of the error (actually accuracy) of your classifier. The accuracy will be measured using the percentage of correct predictions in the test set, i.e. $\frac{1}{n}\sum_{i=1}^{n} I(y_i = \hat{y}_i)$ where I is the indicator function and n is the number of observations in the test set.

Tasks

1. Create the most accurate classifier that you can for the data, as measured by the test data. This will require submitting your predictions to a specially created kaggle site.
2. Write a 3–5 page report summarizing your approach to building the classifier, and whether you've learned anything interesting about the differences in the way Di and Souhaib speak.

Deadlines:

- Sep 23: Create a kaggle account using your monash email address. (If you want to use a different email address we will need to specially invite you to participate.) At least one submission to kaggle. The address for the competition is https://inclass.kaggle.com/c/cockatoo.
- Oct 17: Submissions to kaggle close.
- Oct. 21: Final report due, turned in through moodle.

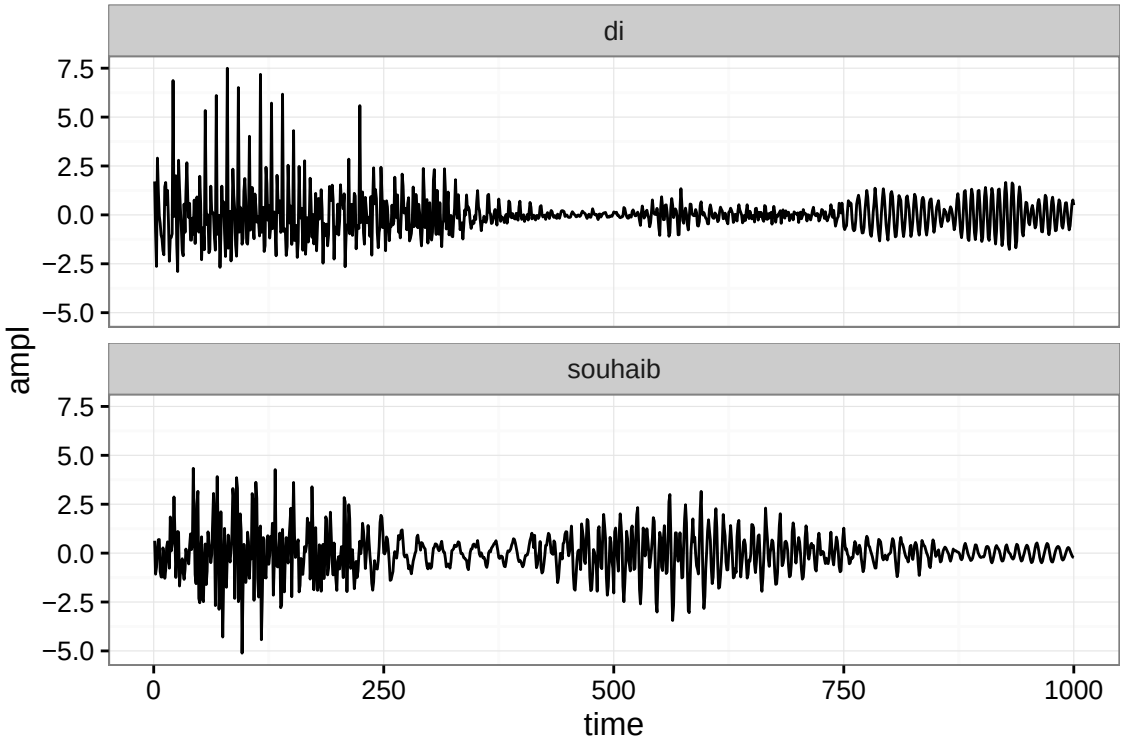
Grading:

- Total points: 20
- Accuracy of classifier: 9 (most accurate scores 9, second most gets 8, third gets 7.5, fourth 7, and then 0.1 less for each successive rank.)
- Report: 9
- Met deadlines: 2

Suggestions

- It might be good to create new variables to help you build a better classifier
- Submit lots of entries to improve your predictions. Submit early and often until the final deadline.

Here is a example of audio signal for the word “bagging” for both Di and Souhaib.



The full audio files look like this: