Lyrics Emotion Detection App

Andrea Ierardi

Student's ID: 960188
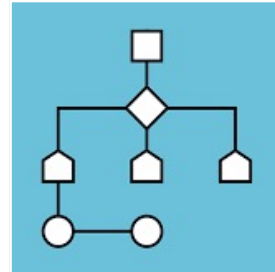
Project aim

Build prediction models to detect emotions in song lyrics.

Build Algorithm for song suggestion

Build a simple Web Application to display the results
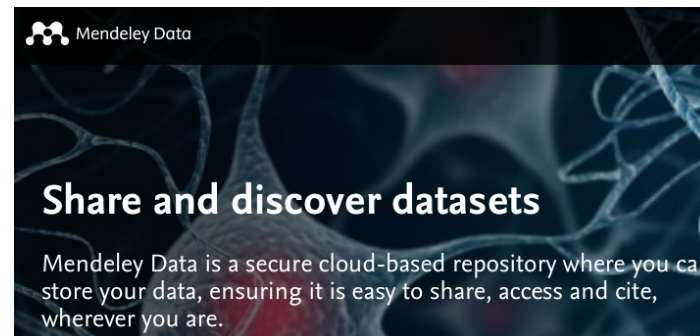
Datasets

- Two datasets chosen:

  - **Training and testing model:**
    - WASSA-2017 SharedTask on Emotion Intensity (EmoInt)

  - **Emotion detection lyrics**:
    - Music Dataset: Lyrics and Metadata from 1950 to 2019

WASSA-2017 Shared Task on Emotion Intensity (EmoInt)

Part of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA-2017), which is to be held in conjunction with EMNLP-2017.

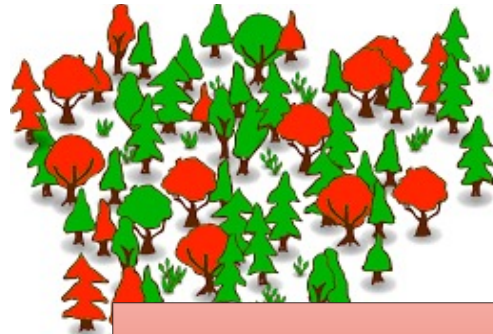*I am interested. How do I get going?*
This task has concluded, but go here for the next iteration: SemEval-2018 Task 1 *Affect in Tweets*.

Mendeley Data

Share and discover datasets

Mendeley Data is a secure cloud-based repository where you can store your data, ensuring it is easy to share, access and cite, wherever you are.
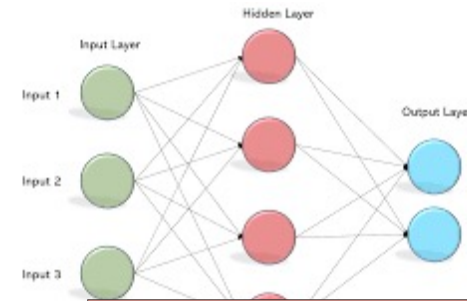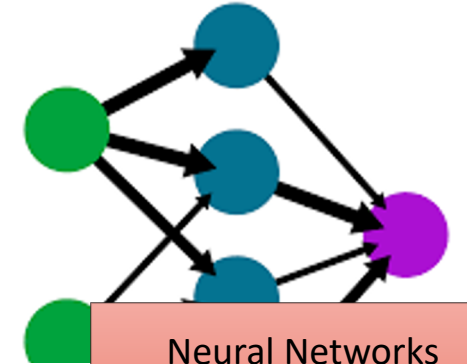
Models

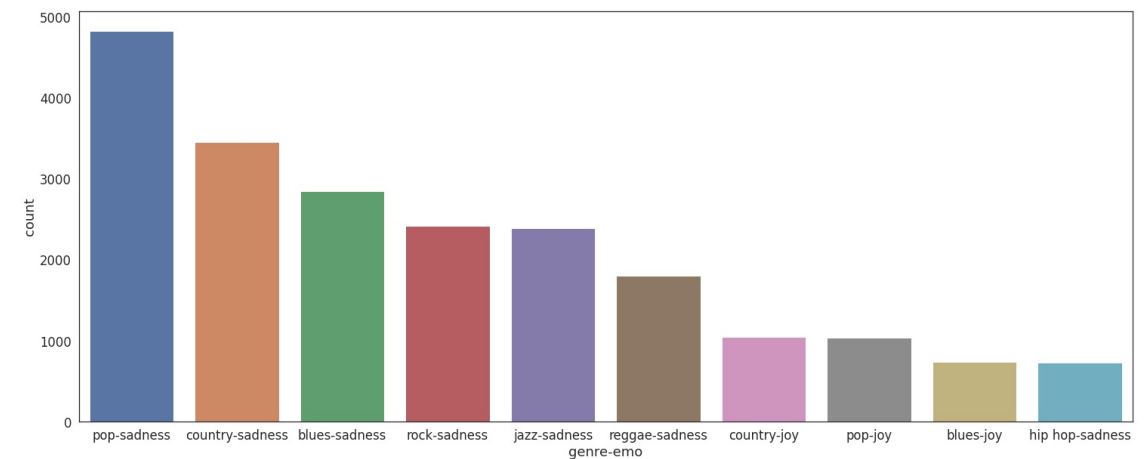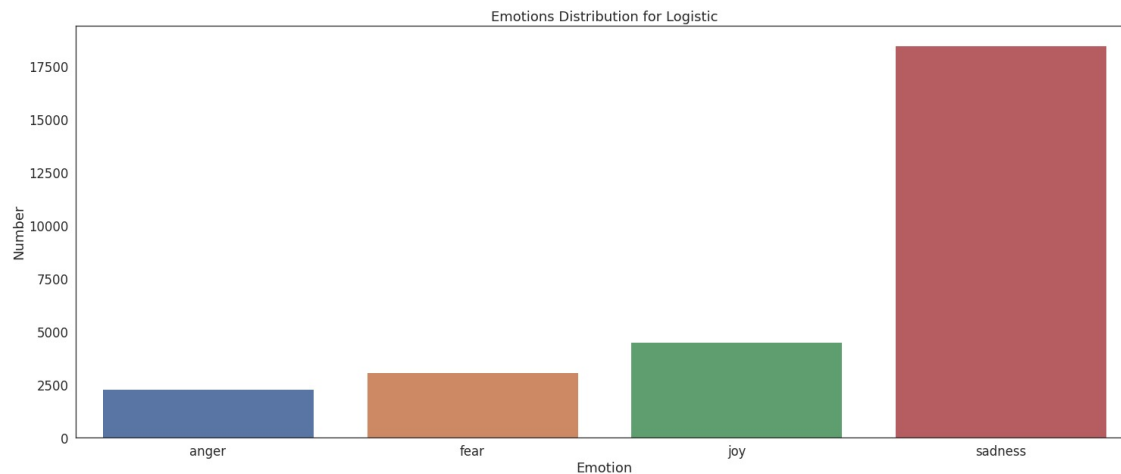Logistic Regression Classifier

Random Forest Classifier

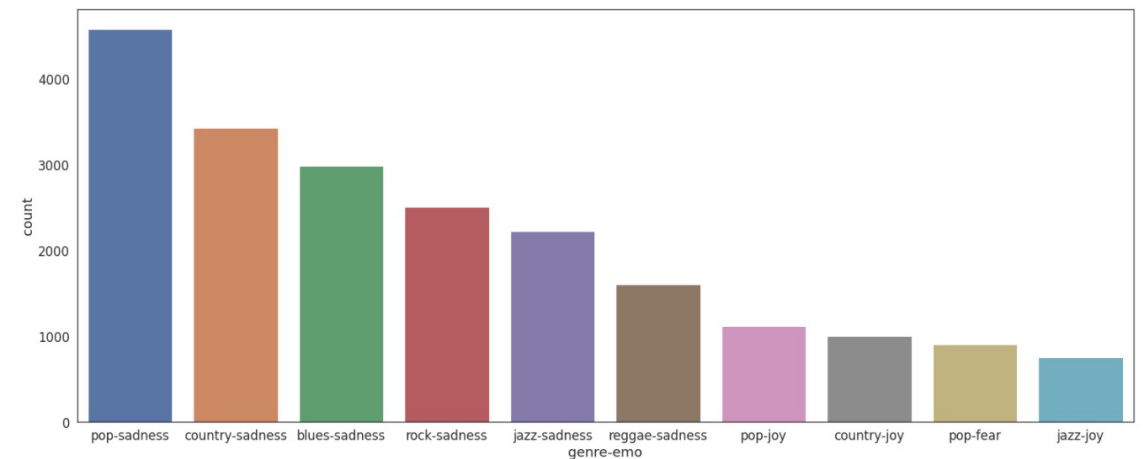Multi-layer Perceptron Classifier

Neural Networks

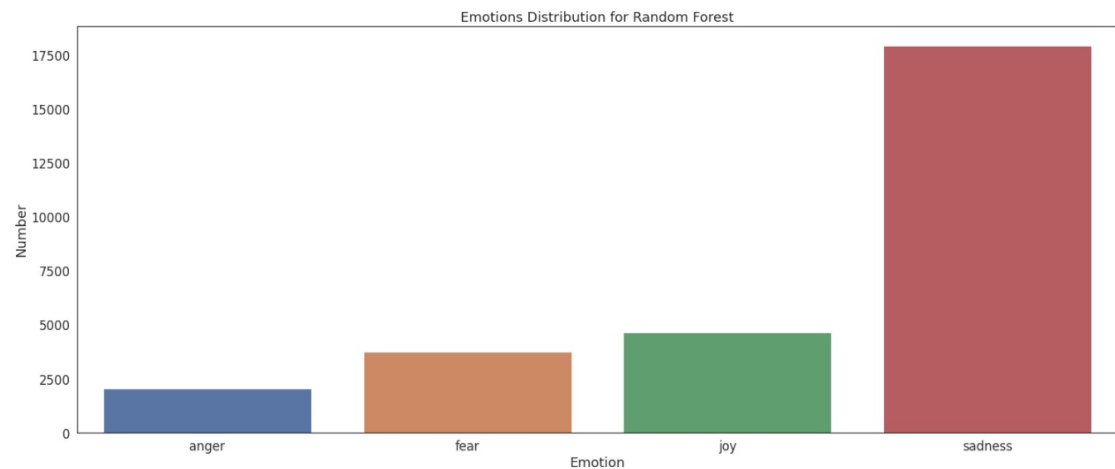# Logistic regression

- F1 score : 0.8263
  - Averaged for multi-class problem

- Most predicted emotion is sadness:
  - It is predicted 60% more than the others
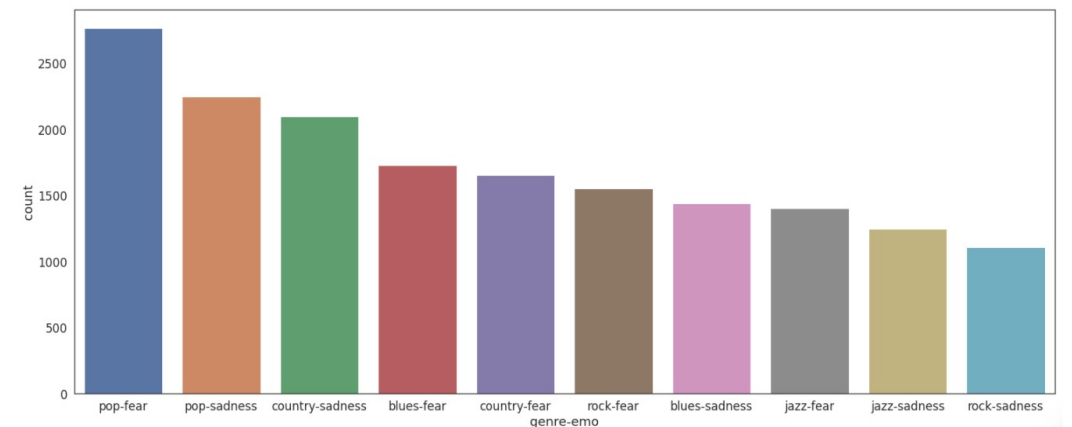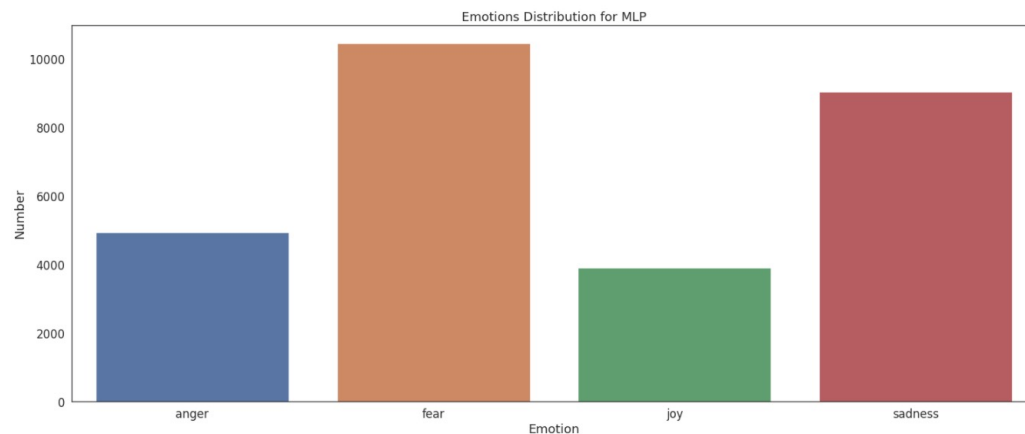

Emotions Distribution for Logistic

Random Forest

- F1 score: 0.802
  - Averaged for multi-class problem

- Most predicted emotion is sadness:
  - It is predicted 60% more than the others

Emotions Distribution for Random Forest

Multi-Layer Perceptron

- F1 score: 0.803
  - Averaged for multi-class problem

- Most predicted emotion is Fear and Sadness

Emotions Distribution for MLP

Neural Networks

Dense Neural Networks:
3 dense layers with different number of nodes.

Neural Networks with Embeddings:
Embedding layer and two Dense layers
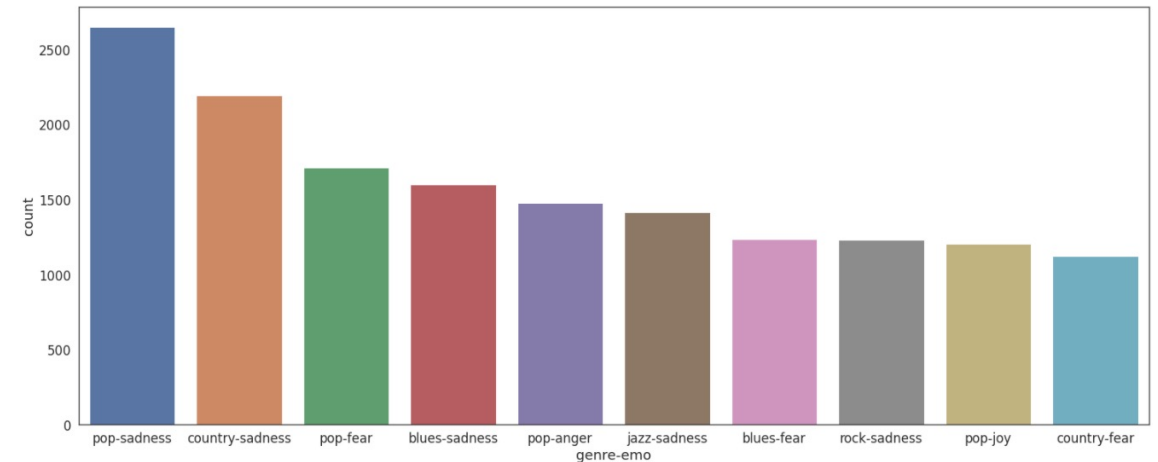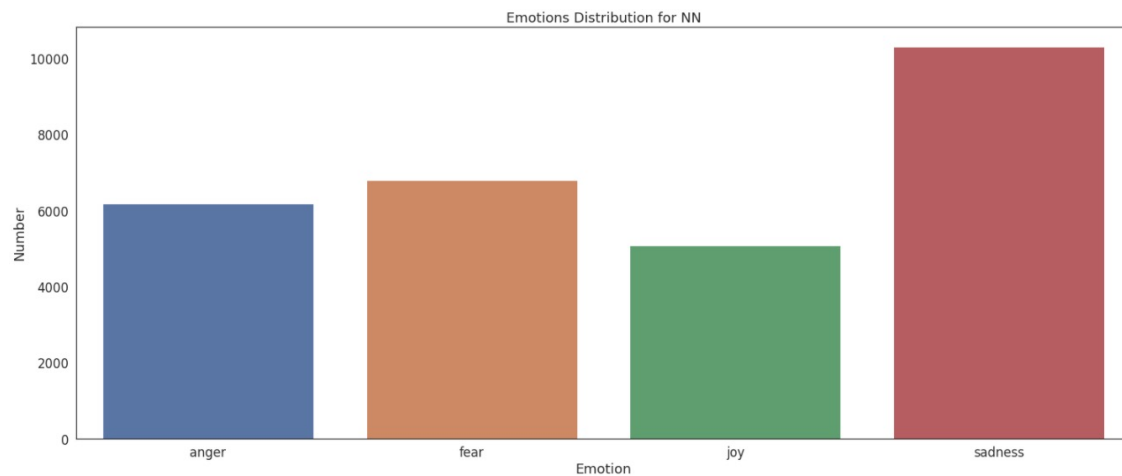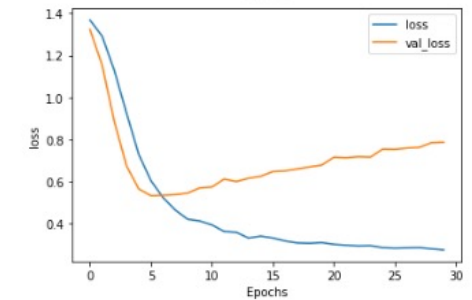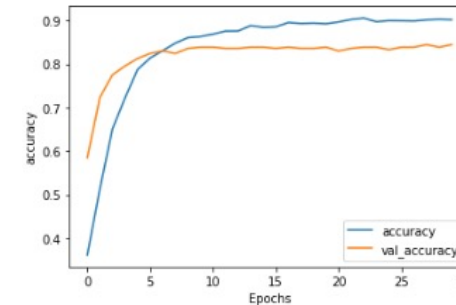
Simple Feed Forward Neural Networks:
1 dense layers

Tuned Dense Neural Networks
with Grid Cross Validation

Best Neural Network result

- Tuned Neural Networks output
  the best f1 score: 0.826

- Most predicted emotion is sadness

- Prediction are more balanced

Emotions Distribution for NN

Song Suggestion

- Simple random song suggestion Algorithm

- **Algorithm:**
  - User input: genre and emotion
  - Retrieve from **Emotion detection lyrics** dataset songs predicted with the same genre and emo as input
  - Return randomly 10 songs

track_name	artist_name	genre	pred
anna lee, the healer	the beach boys	pop	joy
loving is easy	rex orange county	pop	joy
smoke gets in your eyes	the platters	pop	joy
jesus loves me	whitney houston	pop	joy
it's my life	no doubt	pop	joy
4th dimension	kids see ghosts	pop	joy
true believers	the bouncing souls	pop	joy
pay it back	elvis costello	pop	joy
do my thang	miley cyrus	pop	joy
all in the groove	blues traveler	pop	joy

Web Application

- Main page:
  - Form for lyrics test
  - Model prediction results
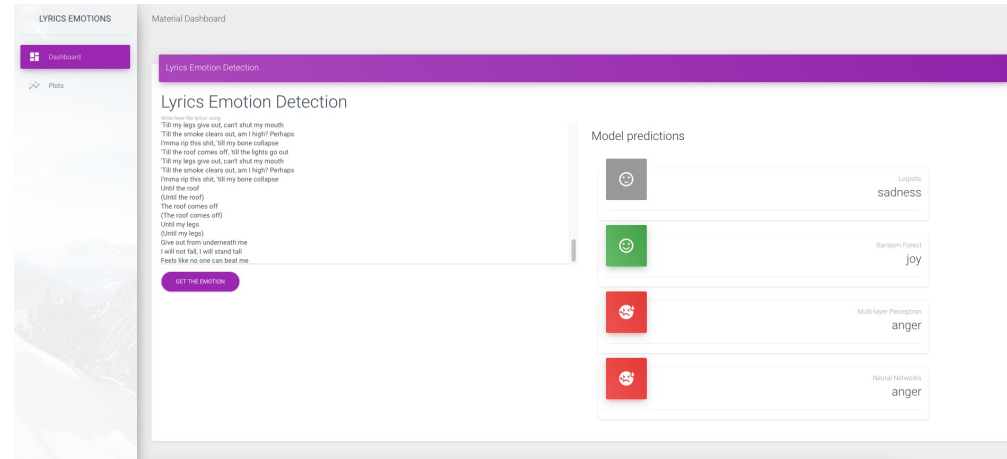
- Plot page:
  - Dataset description
  - Plots

Simple tests

	Logistic
☺	joy

	Random Forest
☹	fear

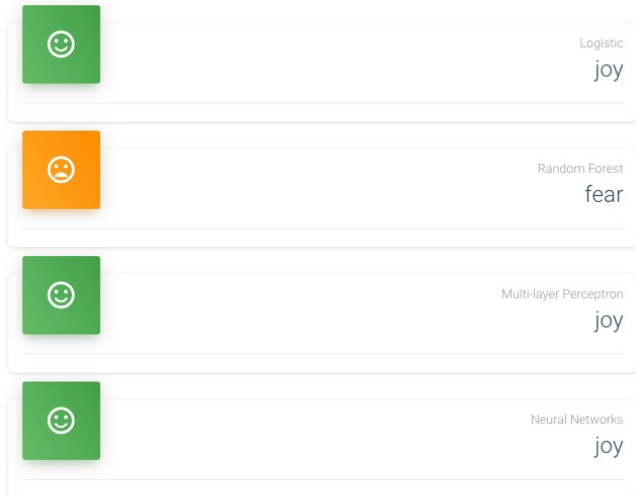|	Multi-layer Perceptron
☺	joy

	Neural Networks
☺	joy

	Logistic
☺	joy

	Random Forest
☹	fear

	Multi-layer Perceptron
☺	sadness

	Neural Networks
☺	joy

	Logistic
☺	joy

	Random Forest
☹	fear

	Multi-layer Perceptron
☺	joy

	Neural Networks
☺	joy
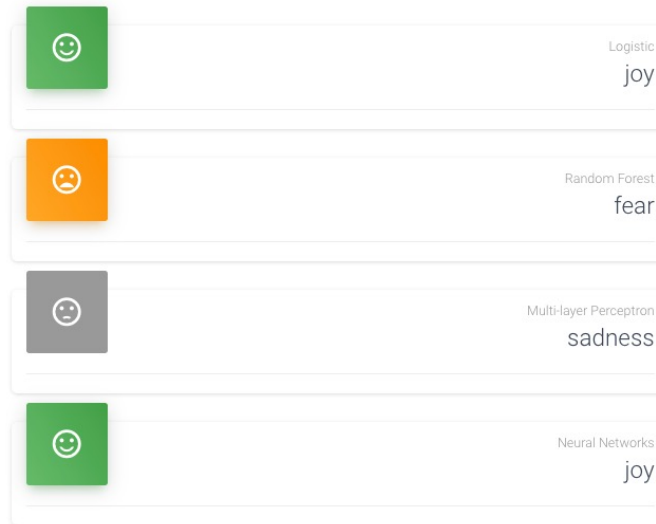
Good day

Bad day

Evil

Real world tests

Model predictions

		Logistic
🙁		sadness

		Random Forest
🙂		sadness

		Multi-layer Perceptron
🙁		sadness

		Neural Networks
🙁		sadness
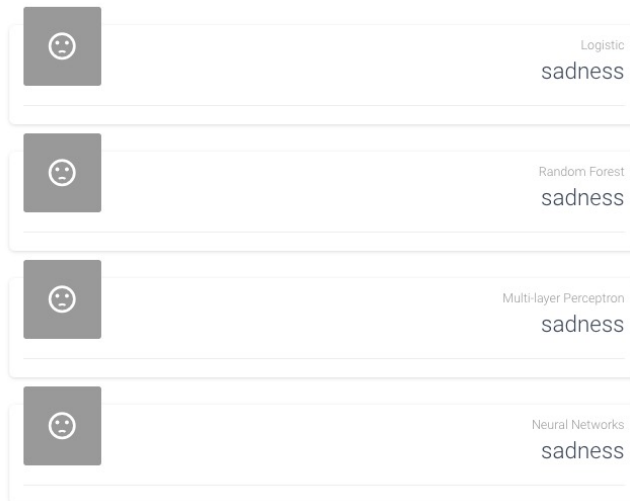
**Blackbirds
The Beatles**

Model predictions

		Logistic
🙂		sadness

		Random Forest
🙂		joy

		Multi-layer Perceptron
🙂		joy

		Neural Networks
🙂		joy
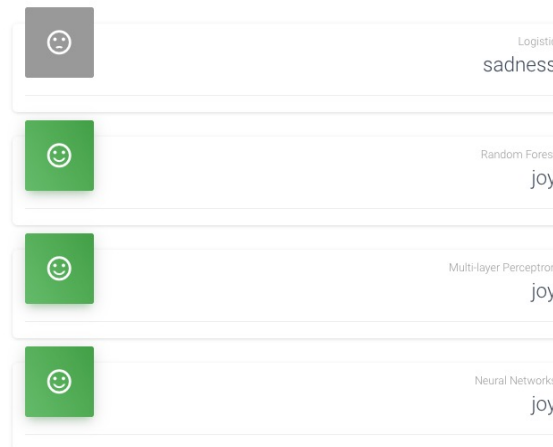
**Good Day Sunshine
The Beatles**

		Logistic
🙁		sadness

		Random Forest
🙂		joy

		Multi-layer Perceptron
😣		anger

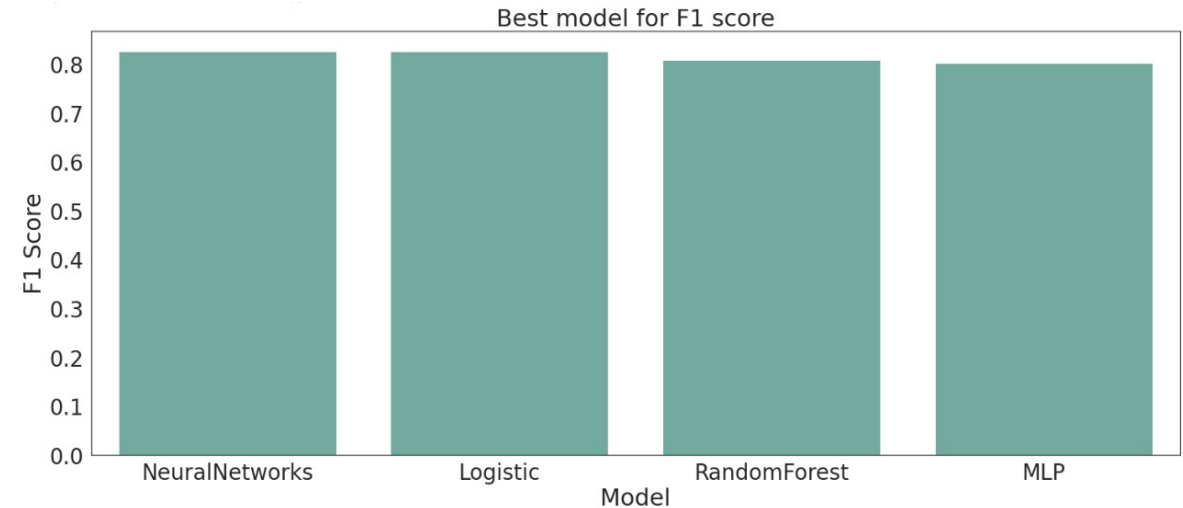|		Neural Networks
😣		anger
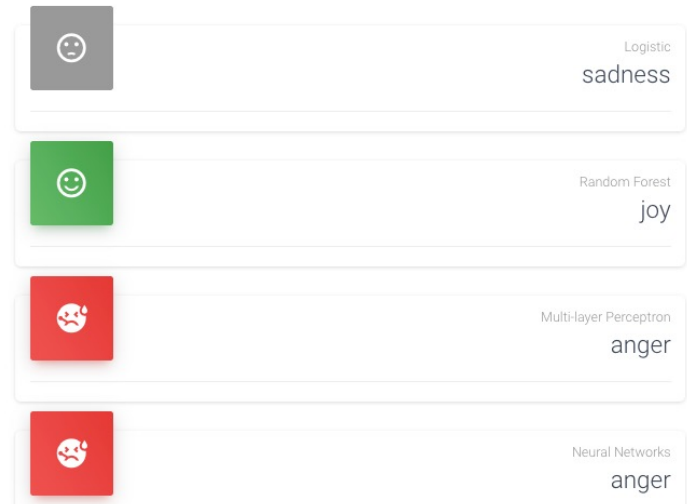
**Till I collapse
Eminem**

Conclusions

- Neural Networks and Logistic outputs similar results according to the F1 score.

- In real world example MLP and Neural Networks outputs better predictions

Questions?