Analyzing Grammatical Facial Expressions

Milestone: Data Collection and Processing

Group 9

Saharsh Desai

Vishal Bothra

857-376-8566 (Tel of Saharsh Desai)

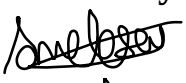
857-269-5186(Tel of Vishal Bothra)

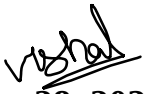
desai.sah@northeastern.edu

bothra.v@northeastern.edu

Percentage of Effort Contributed by Student 1: 50

Percentage of Effort Contributed by Student 2: 50

Signature of Student 1: 

Signature of Student 2: 

Submission Date: February 28, 2022

Project Proposal

IE 7275: Data Mining in Engineering

Problem Setting:
Facial expressions are highly indicative of nonverbal communication cues that play crucial role in interpersonal relations. Primarily we will be focusing on developing a recommender system for playing music based on the mood of the user by comprehending their facial expressions. Challenges that can be faced are capturing the face of the user accurately to predict their mood.

Problem definition:
The goal is to create a recommendation system for popular music streaming platforms like Spotify based on not only user history but also their mood. To do so we will first recognize the mood and classify it into 9 target classes from a dataset generated by using Microsoft Kinect to obtain an image of each frame, identified by a timestamp.

Data Sources:
UCI machine learning repository:
This dataset supports the development of models that make possible to interpret Grammatical Facial Expressions from Brazilian Sign Language (Libras).This dataset was already used in the experiments described in Freitas et al in 2014.
Link: https://archive-beta.ics.uci.edu/ml/datasets/grammatical+facial+expressions

Data Description:

Datapoints files: Coordinates x and y are given in pixels. Coordinates z are given in millimeters.

Feature (x, y, z)	Description	Number of Rows
0-7	left eye	27965
8-15	right eye	27965
16-25	left eyebrow	27965
26-35	right eyebrow	27965
36-47	nose	27965
48-67	mouth	27965
68-86	face contour	27965
87	left iris	27965
88	right iris	27965
89	nose tip	27965
90-94	line above left eyebrow	27965
95-99	line above right eyebrow	27965

Data Collection:

The downloaded data was analyzed to get an overview of the data and decide key metrics for processing the data.

The dataset is constructed from recording sessions with two users. It is composed of eight grammatical markers in Brazilian sign language as mentioned above. These are as follows, along with their respective meaning.

Wh question:

- File naming convention: ‘wh_question_datapoints.txt’ and ‘wh_question_targets.txt’
- Description: It is generally used for questions with ‘When’, ‘Who’, ‘How’, ‘What’, ‘Where’, and ‘Why’
- Example: “What is this?”

Yes/no question:

- File naming convention: ‘yn_question_datapoints.txt’ and ‘yn_question_targets.txt’
- Description: It used when asking a question to which there is a `yes' or `no' answer.
- Example: “Did he go away?”

Doubt question:

- File naming convention: ‘doubts_question_datapoints’ and ‘doubts_question_targets.txt’
- Description: It is a state when an answer is not expected and hence it is not a true question.
- Example: “Did Tony buy A CAR?”

Topics:

- File naming convention: ‘topics_datapoints.txt’ and ‘topics_targets.txt’
- Description: It is used when one of the sentence's parts is displaced to the start of the sentence.
- Example: “I have a pen!”

Negative:

- File naming convention: ‘negative_datapoints.txt’ and ‘negative_targets.txt’
- Description: It is used in negative sentences

- Example: ‘I didn't do that!’

Affirmative:

- File naming convention: ‘affirmative_datapoints.txt’ and ‘affirmative_targets.txt’
- Description: It is used when one affirms for a fact.
- Example: ‘I bought that!’

Conditional:

- File naming convention: ‘conditional_datapoints.txt’ and ‘conditional_targets.txt’
- Description: It used in a subordinate sentence and hence often dependent on the main sentence.
- Example: ‘If you like pizza, you love cheese.’

Emphasis:

- File naming convention: ‘emphasis_datapoints.txt’ and ‘emphasis_targets.txt’
- Description: It is used to highlight new information in the pattern of speech.
- Example: ‘The car is broken.’

Multiple frames were captured from the above-mentioned markers and some predefined points were used to attribute face points.

Data Pre-processing:

For each attribute since its X and Y coordinate points are given as pixel values and the Z coordinate is given in mm. Hence, the numerical range for X, Y differs from that of Z, which makes it essential to standardize Z before using it.

We used standard scaler function from scikit learn to standardize the Z feature associated with all the 100 facial points in the dataset. The standardization will also help in making the model invariant with respect to changes in location of face in captured frame.

We noticed that instead of having NAN values, the dataset consists of 0.0 as a placeholder value. We treated that using imputation by mean as it could result in building an incorrect model.

Each of the marker’s binary classification as conveyed by the target files associated with each marker file, is imbalanced. It means that, it contains an unequal number of positive and negative cases, with on average positive ones being less than the negative ones. Hence to avoid building a model biased towards negative class, we considered equal number of datapoints from both the classes.