

Speech Phoneme Analysis and Classification

By CADET Florent
05/29/2024

Euclidean Distance

K	F1 score \ Run 1	F1 score \ Run 2	F1 score \ Run 3	F1 score \ Run 4	F1 score \ Run 5	F1 score \ Run 6	F1 score \ Run 7	F1 Score Average
2	0.86842	0.86842	0.84211	0.78947	0.86842	0.81579	0.92105	0.85338
3	0.92105	0.97368	0.89474	0.84211	0.92105	0.94737	1.00000	0.92857
4	0.89474	0.94737	0.86842	0.86842	0.89474	0.92105	1.00000	0.91353
5	0.86842	0.94737	0.94737	0.89474	0.92105	0.92105	1.00000	0.92857
6	0.84211	0.97368	0.89474	0.86842	0.89474	0.92105	1.00000	0.91353
7	0.81579	0.97368	0.94737	0.86842	0.89474	0.89474	0.97368	0.90977

Manhattan Distance

K	F1 score \ Run 1	F1 score \ Run 2	F1 score \ Run 3	F1 score \ Run 4	F1 score \ Run 5	F1 score \ Run 6	F1 score \ Run 7	F1 Score Average
2	0.86842	0.86842	0.84211	0.78947	0.89474	0.89474	0.94737	0.87218
3	0.94737	0.94737	0.89474	0.89474	0.94737	0.94737	1.00000	0.93985
4	0.89474	0.94737	0.84211	0.89474	0.89474	0.86842	0.97368	0.90226
5	0.92105	0.97368	0.89474	0.89474	0.89474	0.86842	1.00000	0.92105
6	0.89474	0.97368	0.84211	0.86842	0.89474	0.84211	0.97368	0.89850
7	0.92105	0.97368	0.92105	0.86842	0.89474	0.89474	1.00000	0.92481

Cosine Distance

K	F1 score \ Run 1	F1 score \ Run 2	F1 score \ Run 3	F1 score \ Run 4	F1 score \ Run 5	F1 score \ Run 6	F1 score \ Run 7	F1 Score Average
2	0.89474	0.89474	0.86842	0.89474	0.94737	0.92105	0.94737	0.90977
3	0.97368	0.92105	0.92105	0.89474	0.92105	0.94737	1.00000	0.93985
4	0.97368	0.92105	0.92105	0.86842	0.86842	0.97368	1.00000	0.93233

K	F1 score \ Run 1	F1 score \ Run 2	F1 score \ Run 3	F1 score \ Run 4	F1 score \ Run 5	F1 score \ Run 6	F1 score \ Run 7	F1 Score Average
5	0.97368	0.92105	0.92105	0.89474	0.94737	0.94737	1.00000	0.94361
6	0.94737	0.94737	0.92105	0.86842	0.86842	0.92105	1.00000	0.92481
7	0.97368	0.94737	0.92105	0.86842	0.92105	0.94737	1.00000	0.93985

- **Best Performance:**
 - For **Euclidean Distance**, the highest average F1 score is observed at $K = 3$ and $K = 5$ with 0.92857.
 - For **Manhattan Distance**, the highest average F1 score is observed at $K = 3$ with 0.93985.
 - For **Cosine Distance**, the highest average F1 score is observed at $K = 5$ with 0.94361.

How does performance change with different values of K

- As K increases from 2 to around 5, the average F1 score tends to improve, suggesting better performance with more neighbors considered.
- After $K = 5$, the performance plateaus or slightly decreases, indicating that considering more than 5 neighbors does not significantly improve performance and might even degrade it slightly due to the inclusion of less relevant neighbors.

What distance metric did you use ? Tried any others ?

- **Cosine Distance** consistently shows high performance, with the highest average F1 score of 0.94361 at $K = 5$.
- **Manhattan Distance** also shows high performance, with the highest average F1 score of 0.93985 at $K = 3$.
- **Euclidean Distance** also shows performance but slightly lower compared to Cosine and Manhattan, with the highest average F1 score of 0.92857 at $K = 3$ and $K = 5$.

Optimum classification :

For optimum classification performance, we recommend using the **Cosine Distance** with $K = 5$ based on the highest observed mean F1 score of 0.94361. However, **Manhattan Distance** with $K = 3$ and **Euclidean Distance** with $K = 3$ or $K = 5$ are also viable options with slightly lower but good performance.

How does performance change when classification is done one data for a single gender alone, or when data from both genders are put together?

Using $K = 5$:

Euclidean Distance

Gender	F1 score \ Run 1	F1 score \ Run 2	F1 score \ Run 3	F1 score \ Run 4	F1 score \ Run 5	F1 score \ Run 6	F1 score \ Run 7	F1 Score Average
F	0.94737	0.94737	0.94737	1.00000	1.00000	1.00000	1.00000	0.97744
M	0.89474	0.89474	0.89474	0.84211	0.89474	0.68421	0.78947	0.84211
all	0.86842	0.94737	0.94737	0.89474	0.92105	0.92105	1.00000	0.92857

Manhattan Distance

Gender	F1 score \ Run 1	F1 score \ Run 2	F1 score \ Run 3	F1 score \ Run 4	F1 score \ Run 5	F1 score \ Run 6	F1 score \ Run 7	F1 Score Average
F	0.94737	0.94737	0.94737	1.00000	1.00000	1.00000	1.00000	0.97744
M	0.89474	0.94737	0.84211	0.89474	0.89474	0.73684	0.73684	0.84962
all	0.92105	0.97368	0.89474	0.89474	0.89474	0.86842	1.00000	0.92105

Cosine Distance

Gender	F1 score \ Run 1	F1 score \ Run 2	F1 score \ Run 3	F1 score \ Run 4	F1 score \ Run 5	F1 score \ Run 6	F1 score \ Run 7	F1 Score Average
F	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
M	0.84211	0.94737	0.89474	0.84211	0.78947	0.73684	0.89474	0.84962
all	0.97368	0.92105	0.92105	0.89474	0.94737	0.94737	1.00000	0.94361

The performance of classification varies significantly based on whether data is from a single gender or both genders combined. For Euclidean and Manhattan distances, the F1 scores for females are consistently high, with averages of 0.97744 for both distance measures, indicating robust performance. For males, the performance is lower, with averages of 0.84211 for Euclidean and 0.84962 for Manhattan. When data from both genders are combined, the performance improves, achieving F1 score averages of 0.92857 for Euclidean and 0.92105 for Manhattan. With Cosine distance, females again show perfect performance with an average F1 score of 1.000, while males have a lower average of 0.84962. Combined data results in an average F1 score of 0.94361, indicating that integrating data from both genders generally enhances classification performance, particularly improving the consistency and accuracy compared to using data from males alone.

What are the vowel-based phonemes that produce the most confusion (base this off your confusion matrix)

As seen in the confusion matrices in the folder [conf_metric/](#) the vowel-based phonemes ('AE', 'IY', 'UH') that produce the most confusion, based on the confusion matrices provided, are the phoneme 'UH' and the phoneme 'IY', as they are frequently misclassified relative to each other in all distance measurements (Euclidean, Manhattan, and Cosine). In addition, there is notable confusion between the phoneme 'UH' and the phoneme 'AE', particularly with Manhattan and Cosine distances. This consistent pattern of misclassification underlines the fact that these phoneme pairs share similar characteristics, making them more difficult to distinguish using the given classification methods.