UCLA Samueli School of Engineering

Class Project - Regional Dialect Identification for African American English Speakers

UCLA, Speech Processing and Auditory Perception Lab

1

Background - African American English (AAE)

- AAE is the set of regional English variants commonly spoken by Black people throughout the United States
- Contains phonological (pronunciation), morphosyntactic (grammar + diction), and prosodic differences from Mainstream American English (MAE)

UCLA Samueli School of Engineering

2

Background - African American English (AAE)

Common Examples

Phonological	
Labialization of interdental fricatives	north -> norf
Syllable initial fricative stopping	those -> doze
Metathesis of final /s/+stop	ask -> aks

Morphosyntactic	
Copula or auxiliary verb absence	They gone, I never been there
Y'all/they to mark 2nd personal plural and 3rd plural possessive	It's y'all ball, it's they house
Negative inversion	Can't nobody say nothing

UCLA Samueli School of Engineering

3

3

Background - Dialect Density

- Dialect Density is the amount of dialect usage that appears in one's speech or text
- Dialect Density Measure (DDM)

$$DDM = \frac{\#\ dialectal\ tokens\ in\ utterance}{\#\ words\ in\ utterance}$$

Ex. "Nobody <u>aksed</u> him <u>nothing</u>" -> DDM= 2/4 = 50%

UCLA Samueli School of Engineering

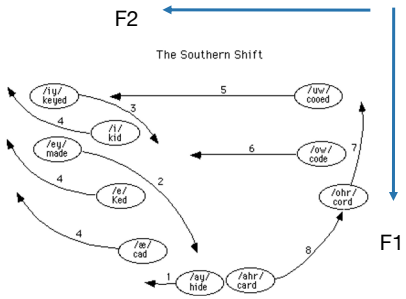
J. A. Washington, L. Branum-Martin, C. Sun, and R. Lee-James, "The impact of dialect density on the growth of language and reading in African American children." Language, speech, and hearing services in schools, 49(2), pp. 232-247. 2018.

4

4

Background - Regional Accent/Dialect

- Speech patterns vary across region along with socioeconomic group and ethnicity
- Eg. The Southern Shift demonstrates how formant values are different between the US South and North



Bailey, Guy, and Erik Thomas. 1998. "Some Aspects of AAVE Phonology." In *African American English: Structure, History, and Use*, edited by Salikoko Mufwene, John Rickford, Guy Bailey, and John Baugh, 85–109. London: Routledge.

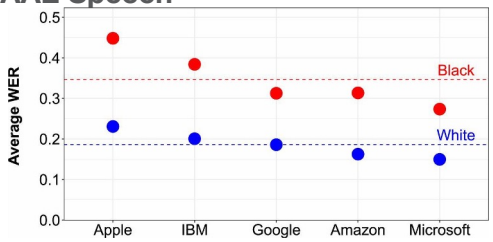
UCLA Samueli School of Engineering

5

5

Background - Relevance to ASR

- ASR systems trained on MAE speech perform much worse for AAE Speech



A.Koenecke, A.Nam, E.Lake, J.Nudell, M.Quartey, Z.Mengesha, C.Toups, J.R. Rickford, D.Jurafsky, and S. Goel, "Racial Disparities in Automated Speech Recognition," *Proceedings of the National Academy of Sciences*, vol. 117, no. 14, pp. 7684–7689, 2020.

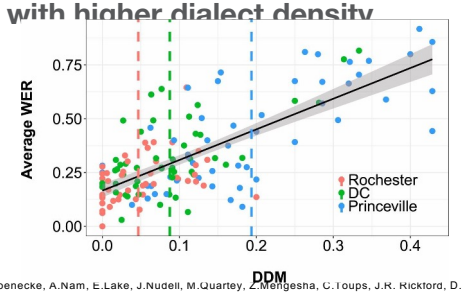
UCLA Samueli School of Engineering

6

6

Background - Relevance to ASR

- This degradation in performance becomes worse for speakers with higher dialect density



A.Koenecke, A.Nam, E.Lake, J.Nudell, M.Quartey, Z.Mengesha, C.Toups, J.R. Rickford, D.Jurafsky, and S. Goel, "Racial Disparities in Automated Speech Recognition," *Proceedings of the National Academy of Sciences*, vol. 117, no. 14, pp. 7684–7689, 2020.

UCLA Samueli School of Engineering

7

7

Motivation

If we can predict a speaker's regional language variant in advance of downstream speech tasks, we can:

- Select an ASR model or model parameters best fit for the user
- Automatically estimate and mitigate bias towards user groups
- Improve speaker recognition and speech biometric tasks

UCLA Samueli School of Engineering

8

8

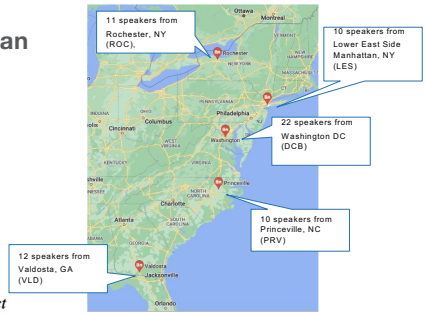
Methods - Dataset

Corpus of Regional African American Language (CORAAL)

- Database of AAE speakers from different US cities

	DDMphon	DDMgram	DDM
DCB	0.083	0.004	0.088
ROC	0.041	0.006	0.047
PRV	0.166	0.028	0.194
LES	0.018	0.025	0.042
VLD	0.122	0.029	0.141

Table 1: *Average dialect density by city for each of the dialect density measures shown.*



T. Kendall and C. Farrington, "The Corpus of Regional African American Language. Version 2021.07." 2021. [Online]. Available: http://oraal.uoregon.edu/coraal

UCLA Samueli School of Engineering

9

9

Methods - Task

Given recordings from speakers across 5 cities, predict the regional dialect of each speaker, you will:

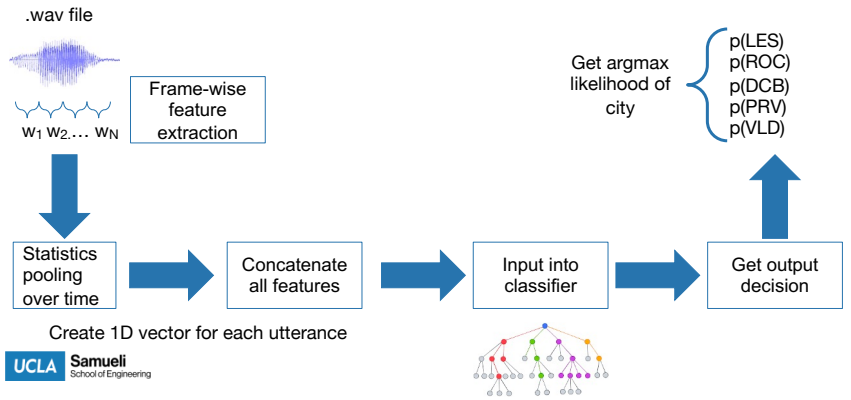
1. **Extract acoustic features related to dialect**
2. **Input the features into the XGBoost classifier**
3. **Measure performance of the system**

UCLA Samueli School of Engineering

10

10

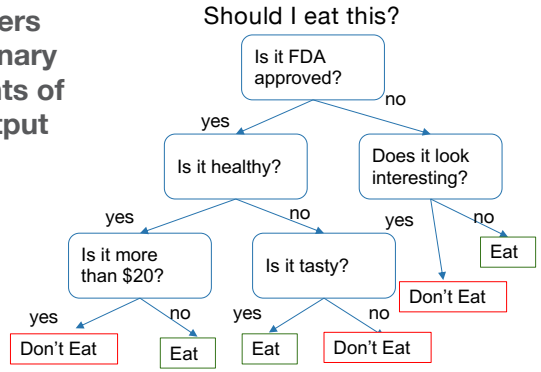
Methods - Possible Pipeline



11

Methods - Decision Tree

Decision trees - classifiers that make sequential binary decisions about elements of the data to make an output decision



12

Methods - XGBoost

Decision tree pros - fast to train and explainable

Decision tree cons - simpler classifier, can't learn overly complex relationships

Extreme Gradient Boosting (XGBoost) is an ensemble method to pool outputs across decision trees to make a more comprehensive decision

UCLA Samueli School of Engineering

13

Methods - XGBoost

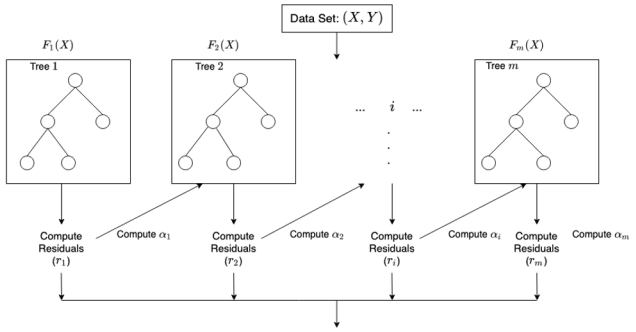
Multiple trees are trained iteratively to accomplish the same task

1. Train a decision tree (Tree 1) to best fit the data
2. Train (Tree 2) to make predictions which correct errors made by (Tree 1)
3. Train (Tree 3) to correct the errors made by mean(Tree 1 + Tree 2)
4. Continue this process until a desired outcome is achieved

UCLA Samueli School of Engineering

14

Methods - XGBoost



$$F_m(X) = F_{m-1}(X) + \alpha_m h_m(X, r_{m-1}),$$

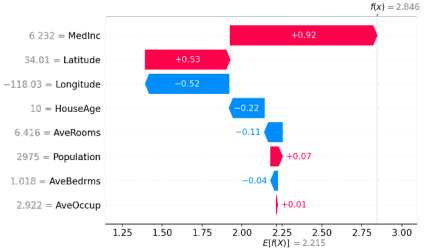
where α_i, and r_i are the regularization parameters and residuals computed with the i^{th} tree respectfully, and h_i is a function that is trained to predict residuals, r_i using X for the i^{th} tree. To compute α_i we use the residuals computed, r_i and compute the following: $arg\min_{\alpha} = \sum_{i=1}^{m} L(Y_i, F_{i-1}(X_i) + \alpha h_i(X_i, r_{i-1}))$ where $L(Y, F(X))$ is a differentiable loss function.

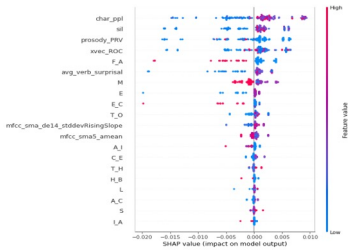
UCLA Samueli School of Engineering

15

Methods - SHAP Value Analysis

Explainability - We can then see how often each input feature was used in the system of decision trees (ie. it's relative contribution in the decision process)





UCLA Samueli School of Engineering

https://shap.readthedocs.io/en/latest/example_notebooks/overviews/An%20introduction%20to%20explainable%20AI%20with%20Shapley%20values.html

16

Challenges

- Robust feature selection
- Combining features in constructive ways
- Dealing with imbalanced data classes
- Creating meaningful representations across time

UCLA Samueli School of Engineering

17