FACULTY OF ARTS AND PHILOSOPHY

GHENT UNIVERSITY – LINGUISTICS DEPARTMENT – MULTIPLES RESEARCH GROUP

Romeo De Timmerman, Stef Slembrouck & Matthias Heyman

INFERENCE VERSUS PREDICTION: INTERPRETING MACHINE LEARNING RESULTS IN (COMPUTATIONAL) SOCIOLINGUISTICS

1 | PREDICTING LANGUAGE USE OF BLUES ARTISTS

This poster builds on a recent publication in *Languages* (cf. QR-code), which examines the **prevalence of African American English features** in the lyrical language use of **blues singers** from various socio-cultural backgrounds and time periods. This study aims to blend insights from the established field of **variationist sociolinguistics**, and the emerging interest in **computational sociolinguistics**. We specifically compiled a **corpus of blues lyrics** and trained a **machine learning algorithm** to **predict** the extent to which blues performers rely on **eight phonological and lexico-grammatical variables of African American English** while singing.

2 | STATISTICS VERSUS MACHINE LEARNING

Most **statistical methods** used in (socio)linguistics are designed to **make inferences** about the **population** based on a **representative sample**. Think of the interpretation of regression coefficients: "the increase in expected value of Y (outcome) for each one-unit increase in X (predictor)". **Machine learning**, on the other hand, focuses on training an algorithm on existing, known data to **predict the outcome in future, unseen data**. Although many such algorithms are seen as "**black-box models**", some are in fact **transparent** (e.g. decision trees can be followed from base to leaf), while others can be **'white-boxed' using post-hoc explanation** methods.

3 | DATA AND METHODS

Data: We compiled a corpus of 270 blues songs (135 originals and 135 covers) performed by 45 artists across three time periods, and three socio-cultural backgrounds.

Analyzed features of AAE

Phonological
- Post-vocalic word-final /r/ deletion
- Post-consonantal word-final /t/ deletion
- Post-consonantal word-final /d/ deletion
- Monophthongization of /aɪ/ diphthongs
- Alveolar nasal /n/ in <ing> ultimas

Lexico-grammatical
- Third person singular <s> deletion
- Not-contraction
- Copula deletion

Analysis: A gradient boosted decision tree model was trained to predict the AAE pronunciation of the eight selected phonological and lexico-grammatical features. The entire dataset comprises 15,184 observations, 70% of which were used for model training, 30% for validation and testing.

Outcome variable
- AAE realization (i.e., whether the AAE feature was realized)

Predictor variables
- Word (i.e., word containing the AAE feature)
- Previous word
- Next word
- Artist name
- Song title
- Song type (i.e., cover or original)
- AAE feature (i.e., one of the eight selected features of AAE)
- Time period (i.e., 1960s, 1980s or 2010s)
- Social group (i.e., AA, non-AA US-based, non-AA non-US-based)

4 | RESULTS AND DISCUSSION

In our entire dataset, artists use the AAE variant of the selected features 74% of the time. Our **trained classification model** was able to **predict this outcome** in the previously unseen test data with an **overall accuracy of 90%** (see paper for detailed model diagnostics). To **white-box the model**, we calculated SHAP values for all predictions. These provide a way to **quantify the contribution of each variable to the model's prediction**, both locally and globally.

Figure 1 shows mean absolute SHAP values on a **global level** for the entire dataset, visualizing the **relative importance of the predictors**. The word containing the AAE feature and the feature itself have the highest impact on model predictions, while **contextual sociolinguistic factors** such as social group and time period **are among the least informative predictors** for the model.

Figure 2 shows SHAP values on a **local level** for a specific datapoint. The mean value, E[f(X)], represents the average prediction, while f(x) indicates the prediction for this specific datapoint. We again conclude that **formal linguistic elements** rather than contextual factors help **explain and predict the use of AAE features by blues artists**. Note that all values are in log(odds) space.

Figure 1: Mean absolute SHAP values for all predictors

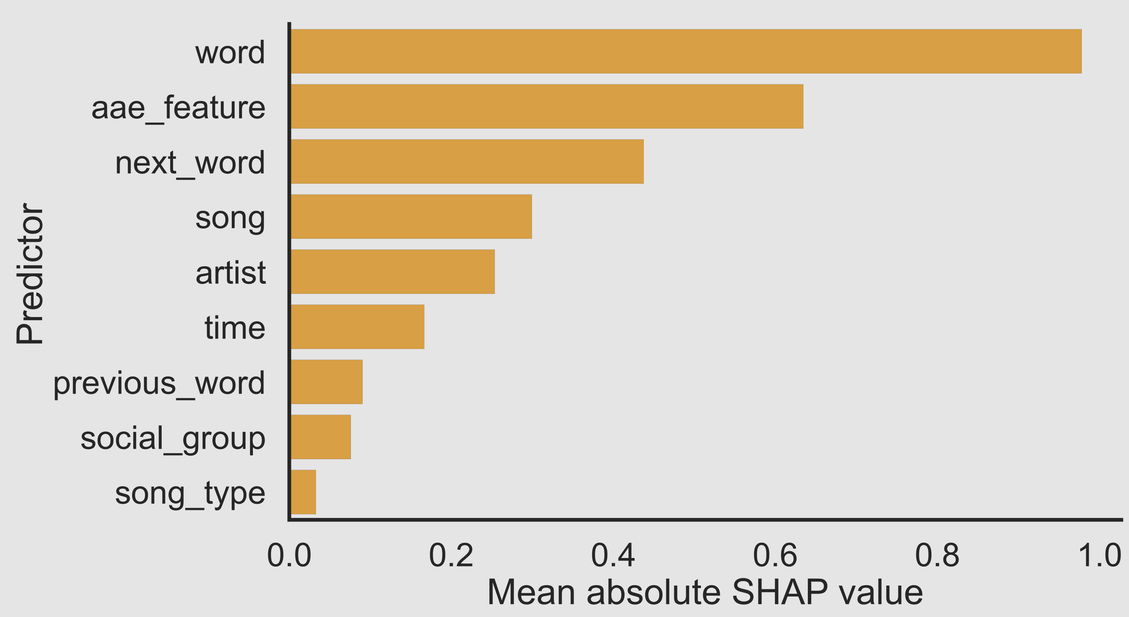
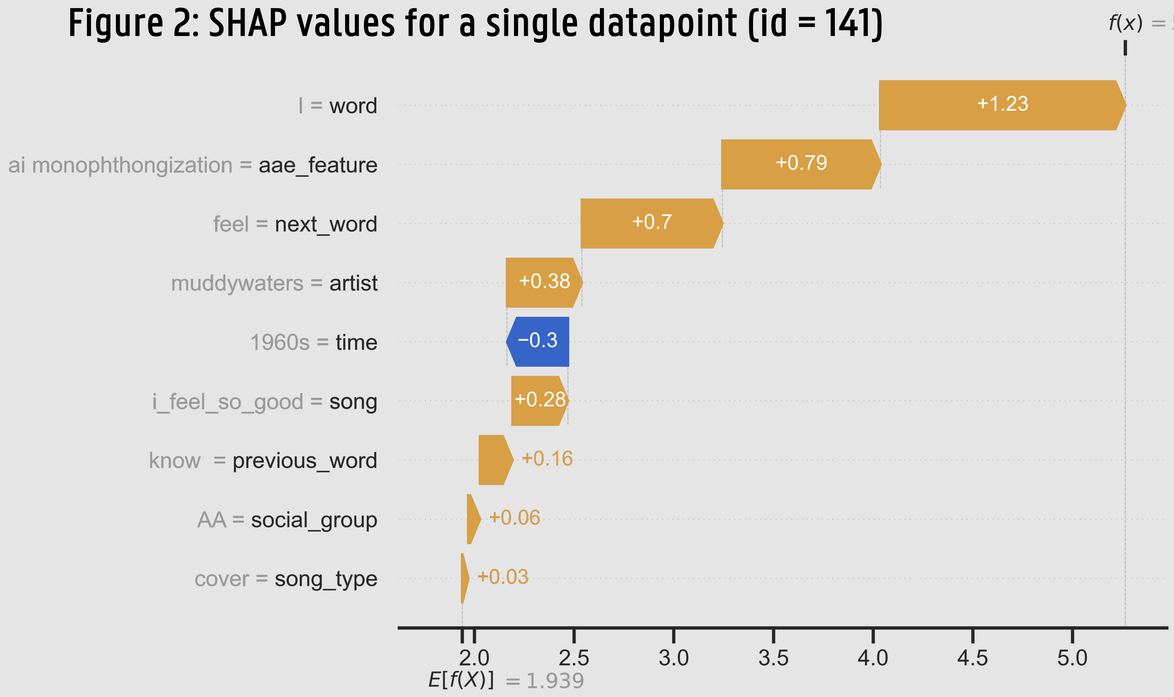


Figure 2: SHAP values for a single datapoint (id = 141)



5 | INTERPRETING MACHINE LEARNING RESULTS

While **many machine learning algorithms** are notoriously **opaque** in terms of their **predictions**, we believe SHAP values and other **model-agnostic explanation methods** can facilitate the **interpretation of quantitative findings into (socio)linguistic insights**. Moreover, we believe that **local predictions and explainability measures** such as those shown in Figure 2 are just **as valuable as broad inferential statements** about a given population. We consequently encourage other scholars in (socio)linguistics to **embrace predictive machine learning methods**.

! Scan QR-code for open access publication !

Questions, feedback, ideas?
- romeo.detimmerman@ugent.be
- romeodetimmerman.dev

GHENT UNIVERSITY