LIST OF FIGURES

Figure 1. Spoken term detection can be partitioned into two tasks: indexing and search. One common approach to indexing is to use a speech to text system (after Fiscus et al., 2007).

Figure 2. A prototype of a web-based application that predicts voice keyword search term reliability is shown. The search term reliability is automatically updated as the user types a search term. A demonstration is available at *http://www.isip.piconepress.com/projects/ks\_prediction/demo/current/*.

Figure 3. In our approach to predicting search term reliability, we decompose terms into features, such as N‑grams of phonemes and the number of phonemes, and apply these features to a variety of machine-learning algorithms.

Figure 4. The relationship between duration and error rate shows that longer words generally result in better performance, but the overall variance of this measure is high.

Figure 5. Feature importance based on the RF algorithm is shown. The feature ”count,” which represents the frequency of occurrence of a word, is by far the singlemost valuable feature since it is not correlated with any of the other features.

Figure 6. The predicted error rate is plotted against the reference error rate, demonstrating good correlation between the two.

# Figures



Figure 1. Spoken term detection can be partitioned into two tasks: indexing and search. One common approach to indexing is to use a speech to text system (after Fiscus et al., 2007).



Figure 2. A prototype of a web-based application that predicts voice keyword search term reliability is shown. The search term reliability is automatically updated as the user types a search term. A demonstration is available at *http://www.isip.piconepress.com/projects/ks\_prediction/demo/current/*.



Figure 3. In our approach to predicting search term reliability, we decompose terms into features, such as N‑grams of phonemes and the number of phonemes, and apply these features to a variety of machine-learning algorithms.



Figure 4. The relationship between duration and error rate shows that longer words generally result in better performance, but the overall variance of this measure is high.



Figure 5. Feature importance based on the RF algorithm is shown. The feature ”count,” which represents the frequency of occurrence of a word, is by far the singlemost valuable feature since it is not correlated with any of the other features.



Figure 6. The predicted error rate is plotted against the reference error rate, demonstrating good correlation between the two.