LIST OF TABLES

# Tables

|  |  |  |  |
| --- | --- | --- | --- |
| **Method** | **No. Feats** | **MSE (Train)** | **MSE (Eval)** |
| All Features / LR/ Corr | 150 | 0.015 | 0.018 |
| SFS / LR / Corr | 55 | 0.016 | 0.017 |
| SFS / LR / MSE | 54 | 0.016 | 0.017 |
| SFS / NN / Corr | 12 | 0.015 | 0.015 |
| SFS / NN / MSE | 14 | 0.015 | 0.015 |
| SFS / Tree / Corr | 7 | 0.015 | 0.020 |
| SFS / Tree / MSE | 7 | 0.016 | 0.019 |
| RF | 56 | 0.006 | 0.014 |

Table . The number of features is shown for different feature selection methods as a function of the mean square error (MSE) on both the training and test sets. Performance for the correlation and MSE criteria was comparable.

|  |  |
| --- | --- |
| **Class** | **Phonemes** |
| Silence | sp sil |
| Stops | b p d t g k |
| Fricatives | jh ch sh s z zh f th v dh hh |
| Nasals | m n ng en |
| Liquids | l el r w y |
| Vowels | iy ih eh ey ae aa aw ay ah ao ax oy ow uh iw er |

Table . A mapping of phones to broad phonetic classes is shown.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | | **Train** | | **Eval** | |
| **Set** | **K** | **k** | **MSE** | **R** | **MSE** | **R** |
| 1 | 1 | 1 | 0.027 | 0.227 | 0.027 | 0.270 |
| 1 | 1 | 3 | 0.025 | 0.340 | 0.025 | 0.370 |
| 1 | 1 | 5 | 0.024 | 0.394 | 0.023 | 0.425 |
| 1 | 1 | 30 | 0.021 | 0.528 | 0.020 | 0.543 |
| 1 | 1 | inf | 0.023 | 0.456 | 0.022 | 0.471 |
| 1 | 2 | 1 | 0.026 | 0.293 | 0.025 | 0.330 |
| 1 | 2 | 3 | 0.024 | 0.414 | 0.023 | 0.444 |
| 1 | 2 | 5 | 0.022 | 0.461 | 0.022 | 0.473 |
| 1 | 2 | 30 | 0.019 | 0.569 | 0.019 | 0.583 |
| 1 | 2 | inf | 0.018 | 0.601 | 0.018 | 0.615 |
| 1 | 3 | 5 | 0.022 | 0.475 | 0.022 | 0.497 |
| 1 | 3 | 30 | 0.019 | 0.565 | 0.019 | 0.579 |
| 1 | 3 | inf | 0.018 | 0.600 | 0.018 | 0.614 |
| 1 | 4 | 5 | 0.022 | 0.477 | 0.021 | 0.499 |
| 1 | 4 | 30 | 0.020 | 0.542 | 0.020 | 0.559 |
| 1 | 4 | inf | 0.019 | 0.578 | 0.018 | 0.595 |
| 1 | 12 | 5 | 0.024 | 0.397 | 0.023 | 0.432 |
| 1 | 12 | 30 | 0.021 | 0.503 | 0.021 | 0.520 |
| 1 | 12 | inf | 0.021 | 0.519 | 0.020 | 0.542 |
| 2 | 2 | 5 | 0.024 | 0.387 | 0.024 | 0.407 |
| 2 | 4 | inf | 0.020 | 0.550 | 0.019 | 0.568 |
| 2 | 15 | inf | 0.021 | 0.526 | 0.020 | 0.551 |
| 2 | 17 | inf | 0.021 | 0.526 | 0.020 | 0.551 |

Table . The correlation of predicted error rates with actual error rates is shown for our acoustic distance measure. Performance on the eval set is comparable for sets 1 and 2 for a broad range of parameter settings. The correlation between open set and closed set performance is also good.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Train** | | **Eval** | |
| **k** | **MSE** | **R** | **MSE** | **R** |
| 1 | 0.026 | 0.296 | 0.026 | 0.322 |
| 3 | 0.024 | 0.405 | 0.024 | 0.421 |
| 5 | 0.023 | 0.434 | 0.023 | 0.451 |
| 30 | 0.021 | 0.502 | 0.021 | 0.519 |
| 50 | 0.021 | 0.503 | 0.021 | 0.519 |
| 100 | 0.021 | 0.499 | 0.021 | 0.515 |
| 300 | 0.022 | 0.483 | 0.022 | 0.498 |
| inf | 0.023 | 0.459 | 0.022 | 0.478 |

Table . Results are shown for the phonetic distance algorithm as a function of the number of nearest neighbors used in kNN.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Classifier** | **LR** | | **NN** | | **RF** | |
| **Method** | **Train** | **Eval** | **Train** | **Eval** | **Train** | **Eval** |
| All Features / LR/ Corr | 0.683 | 0.618 | 0.724 | 0.624 | 0.895 | 0.708 |
| SFS / LR / Corr | 0.654 | 0.629 | 0.753 | 0.692 | 0.875 | 0.701 |
| SFS / LR / MSE | 0.654 | 0.629 | 0.735 | 0.686 | 0.857 | 0.697 |
| SFS / NN / Corr | 0.571 | 0.573 | 0.697 | 0.691 | 0.776 | 0.676 |
| SFS / NN / MSE | 0.573 | 0.574 | 0.697 | 0.689 | 0.799 | 0.679 |
| SFS / Tree / Corr | 0.561 | 0.564 | 0.674 | 0.669 | 0.761 | 0.659 |
| SFS / Tree / MSE | 0.561 | 0.564 | 0.674 | 0.669 | 0.761 | 0.659 |
| RF | 0.635 | 0.604 | 0.734 | 0.675 | 0.882 | 0.703 |

Table . A comparison of the different classification algorithms as a function of the feature sets is shown. R values are shown (the MSE results follow the same trend). Random forests (RF) give very stable results across a wide range of conditions.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Train** | | **Eval** | | **Relative Contribution** | | |
| **Machines** | **MSE** | **R** | **MSE** | **R** | **Acoustic** | **Phonetic** | **Feature** |
| All | 0.00092 | 0.913 | 0.012 | 0.760 | 41.1% | 10.5% | 48.3% |
| NN+RF | 0.00084 | 0.918 | 0.012 | 0.762 | 44.7% | 15.7% | 39.5% |

Table . Performance improves slightly by combining many predictors using PSO. The acoustic and feature-based metrics contribute equally to the overall result.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Acoustic** | **Phonetic** | **Feature** |
| **Acoustic** | 1 | 0.4 | 0.6 |
| **Phonetic** | 0.4 | 1 | 0.7 |
| **Feature** | 0.6 | 0.7 | 1 |

Table . The correlation between various classifiers is shown. The acoustic-based distance is least correlated with the phonetic-based approach, indicating there could be a benefit to combining these predictors.