

Table 1: Performance of the Multinomial naive Bayes classifier after applying different feature selection approaches on the 200-song validation dataset. . The true positive rate was calculated from songs labeled as happy that were correctly classified, and the false positive rate was calculated from sad songs that were misclassified as happy. ACC = accuracy, PRE = precision, REC = recall, F1 = F1-score, ROC = receiver operating characteristic area under the curve, wl = white list, porter = Porter Stemming, Tf = term frequencies, Tf = term-frequency-inverse document frequencies. The final model is denoted with an asterisk.

	ACC	PRE	REC	F1	ROC
BNB	70.50	75.00	65.71	70.05	70.75
BNB porter	70.00	74.19	65.71	69.70	70.23
BNB wl	65.50	65.25	73.33	69.06	65.09
BNB porter+wl	67.00	66.96	73.33	70.00	66.67
MNB Tf*	72.00*	78.82*	63.81*	70.53*	72.43*
MNB Tf porter	68.00	75.95	57.14	65.22	68.57
MNB Tf wl	64.50	66.67	64.76	65.70	64.49
MNB Tf porter+wl	63.50	66.00	62.86	64.39	63.53
MNB TfidfVec	59.50	80.00	30.48	44.14	61.03
MNB TfidfVec porter	60.50	84.21	30.48	44.76	62.08
MNB TfidfVec wl	65.50	70.45	59.05	64.25	65.84
MNB TfidfVec porter+wl	63.50	68.60	56.19	61.78	63.88

Fact sheet:

- trained on 1000-song training dataset, tested on 200-song validation dataset (positive class: *happy*, negative class: *sad*)
- stop word removal, porter stemming, and *white list* of positive and negative terms based on [1]
- 1-gram *bag of words* model based on term frequencies (tf) or term frequency-inverse document frequencies (tfidf),
- uniform priors, Laplace smoothing parameter $\alpha = 1.0$

[1] M. Hu and B. Liu. Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pages 168-177. ACM, 2004.

Table 2: Performance of the Multinomial naive Bayes classifier after applying different feature selection approaches on the 1000-song training dataset. . The true positive rate was calculated from songs labeled as happy that were correctly classified, and the false positive rate was calculated from sad songs that were misclassified as happy. ACC = accuracy, PRE = precision, REC = recall, F1 = F1-score, ROC = receiver operating characteristic area under the curve, wl = white list, porter = Porter Stemming, Tf = term frequencies, Tf = term-frequency-inverse document frequencies.

	ACC	PRE	REC	F1	ROC
BNB	96.50	93.82	98.65	96.17	96.71
BNB porter	93.40	89.26	96.86	92.90	93.74
BNB wl	81.10	72.12	93.95	81.60	82.35
BNB porter+wl	78.80	69.97	91.93	79.46	80.08
MNB Tf	95.70	94.09	96.41	95.24	95.77
MNB Tf porter	93.60	92.83	92.83	92.83	93.52
MNB Tf wl	81.20	75.80	84.98	80.13	81.57
MNB Tf porter+wl	80.50	75.88	82.51	79.05	80.70
MNB TfIdfVec	90.50	99.72	78.92	88.11	89.37
MNB TfIdfVec porter	86.20	99.04	69.73	81.84	84.59
MNB TfIdfVec wl	84.80	81.96	84.53	83.22	84.77
MNB TfIdfVec porter+wl	83.20	80.48	82.29	81.37	83.11