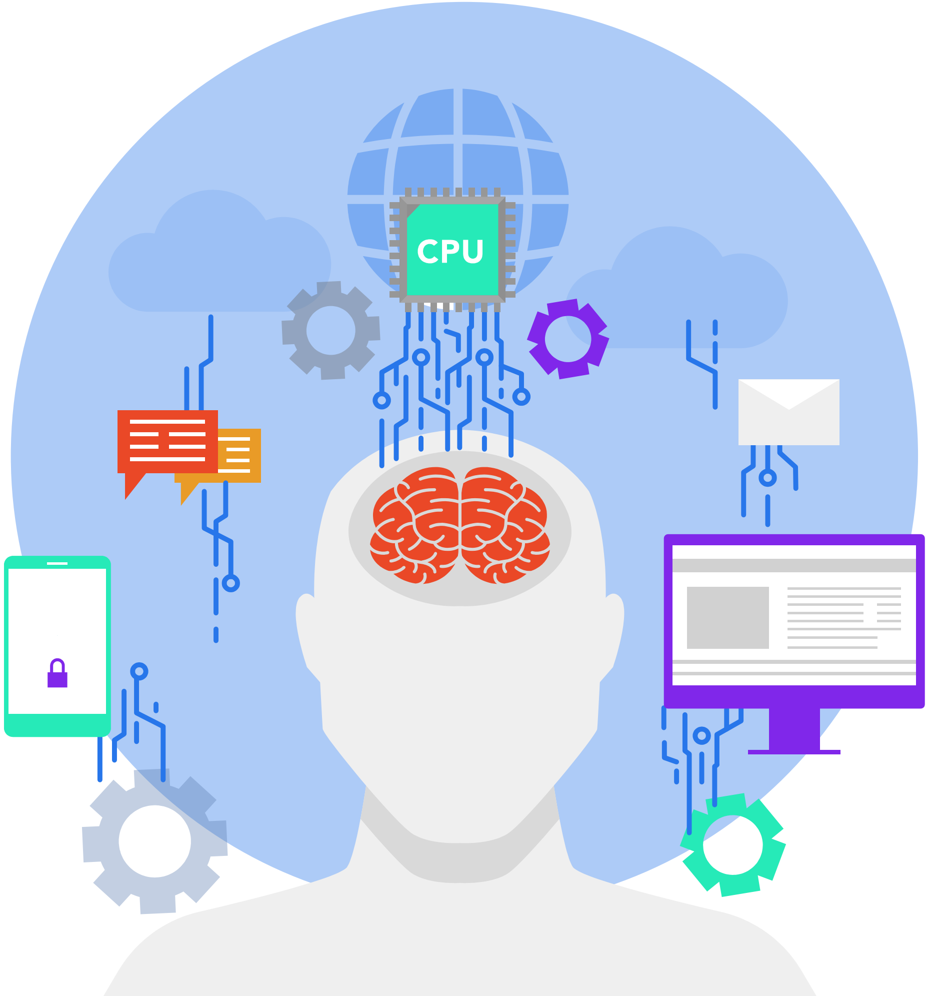
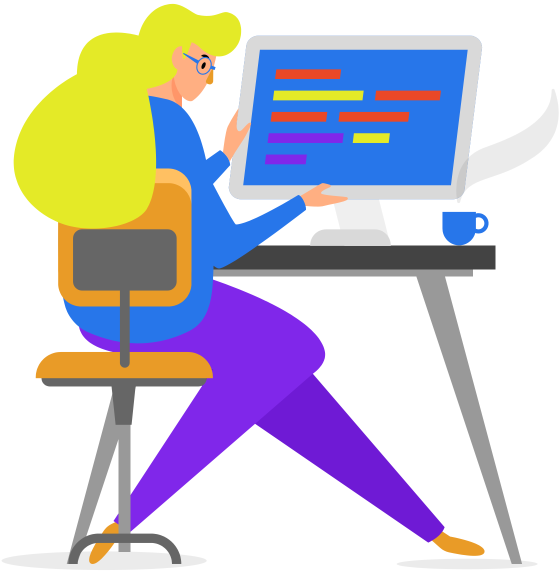


Weighted Naive Bayes

Eshwar Moorthy,
John Kim

Naive Bayes Summary



01 Input Data

Show each instance as a vector

02 Calculate Probabilities

Compute the probabilities per class

03 Compute Each Features' Likelihoods

Compute probabilities for attribute values

04 Independence Assumption

Treat features as conditionally independent

05 Apply Bayes' Theorem

Combine prior and likelihood probabilities

06 Classify Based on Maximum Probability

Choose class with the highest probability

Independent vs Dependent

Independent events

- Outcomes don’t affect each other
- $P(A \cap B) = P(A) \cdot P(B)$
- Probabilities remain constant

Vs

Dependent events

- Outcomes do affect each other
- $P(A \cap B) = P(A) \cdot P(B \mid A)$
- One probability affects the other

Naive Bayes Independence Assumption

Naive Bayes

makes the assumption that **all features are independent of each other** for easier calculations

Prior Methods to Tackle Dependent Features in Naive Bayes

01 TAN

Tree-Augmented Naive Bayes

LBR 02

Lazy Bayesian Rules

03 CLL

Conditional Log Likelihood Optimization

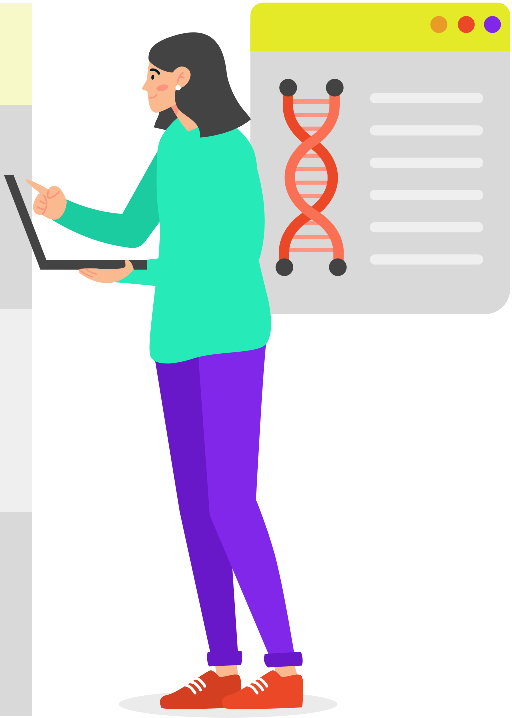
MSE 04

Mean Squared Error Optimization

Datasets

What is Conditional Mutual Information (CMI)?

Datasets	Description
Balance-scale	4 features The average CMI is 0
Cancer	10 features The average CMI is 0.06
Vote	17 features The average CMI is 0.11



Methods

Default Naive Bayes

$$P(class = y \mid X) = P(class = y) * \prod_{i=1}^{\#\,of\,features} P(X_i \mid class = y) \quad (1)$$

$$P(class = y \mid X;\, w) = \gamma_{yx}(w) = P(class = y) * \prod_{i=1}^{\#\,of\,features} P(X_i \mid class = y)^{w_i} \quad (2)$$

$$LCP(w) = \sum_{i=1}^{\#\,of\,features} log\left(\frac{\gamma_{yx}(w)}{\sum_{y'=1}^{\#\,of\,labels} \gamma_{y'x}(w)}\right) \quad (3)$$

Logarithmic Conditional Probability

Weighted Naive Bayes

Results and Analysis

Datasets		
Balance-scale **Average CMI: 0**	Default Naive Bayes: **90.4%** Weighted Naive Bayes: **90.4%** Decision Tree Classifier: **78.4%**	+0%
Cancer **Average CMI: 0.06**	Default Naive Bayes: **72.41%** Weighted Naive Bayes: **75.86%** Decision Tree Classifier: **65.52%**	+3%
Vote **Average CMI: 0.11**	Default Naive Bayes: **89.66%** Weighted Naive Bayes: **94.25%** Decision Tree Classifier: **93.1%**	+4.5%

Conclusion

Modified Naive Bayes

- Accuracies were higher
- Model's predictive power increased
- Dependency was accounted for

Vs

Applications/Further Research

- Can be used during datasets with some features requiring user input
    - Dependency with those features
- Hybrid models working with large size datasets
- Minimize run time

Source Citations

Alleviating naive Bayes attribute independence assumption ... (n.d.). https://jmlr.org/papers/volume14/zaidi13a/zaidi13a.pdf

Shushrut. (2019, April 28). *Balance scale*. Kaggle. https://www.kaggle.com/datasets/mysticvalley/balance-scale

Webb, G. I., Boughton, J. R., & Wang, Z. (2017, July 21). *Not so naive bayes: Aggregating one-dependence estimators - machine learning*. SpringerLink. https://link.springer.com/article/10.1007/s10994-005-4258-6

scipy for backpropagation to get optimal weights and scikit-learn for Decision Tree classifier