Thinkful Data Science Supervised Learning Algorithms and Use Cases

By Arun Nair

Unit 3 – Identifying Models to use challenge

You now have a fairly substantial starting toolbox of supervised learning methods that you can use to tackle a host of exciting problems. To make sure all of these ideas are organized in your mind, please go through the list of problems below. For each, identify which supervised learning method(s) would be best for addressing that particular problem. Explain your reasoning and discuss your answers with your mentor.

1. Predict the running times of prospective Olympic sprinters using data from the last 20 Olympics.

   Since the data here is continuous, and we are looking to find an amount (time), I would choose different regression algorithms. K-NN, Ridge.

   https://ieeexplore.ieee.org/document/7813147

2. You have more features (columns) than rows in your dataset.

   Use a dimension reduction technique such as PCA.

3. Identify the most important characteristic predicting likelihood of being jailed before age 20.

   First do logistic regression. Depending on outcome, create an L1 regularized regression.

4. Implement a filter to "highlight" emails that might be important to the recipient

   Use Naïve Bayes Classifier.

5. You have 1000+ features.

   If the columns are in the thousands, recheck the data and be sure to reevaluate the hypothesis.

6. Predict whether someone who adds items to their cart on a website will purchase the items.

7. K-NN to classify groups of shoppers. Decision Tree Classification to predict yes or no. (Collaborative Filtering).

8. https://towardsdatascience.com/use-algorithms-to-recommend-items-to-customers-in-python-347b769b21f3

9. Your dataset dimensions are 982400 x 500

   Dimension Reduction Techniques (Ridge, Lasso)

10. Identify faces in an image.

    Classification (Decision tree)

11. Predict which of three flavors of ice cream will be most popular with boys vs girls.

    Multivariate/Binary Logistic Regression.

https://medium.freecodecamp.org/using-machine-learning-to-predict-the-quality-of-wines-9e2e13d7480d