Abstract

Index:

1. Introduction
2. Statistical Analysis
   1. Dataset summary
   2. Binomial and Poisson Distribution
   3. Normal Distribution
   4. Importance of Binomial and Poisson and variable justification. Discrete variables in a Normal Distribution
3. Data Preparation and Visualization
   1. EDA
   2. Method justification
   3. Visualisations
   4. Tuft Principles
4. ML
   1. Project management framework CRISP-DM, KDD and SEMMA. unsupervised, or semi-supervised machine learning technique.
   2. ML model.
   3. Results of ML models selected
   4. Similarities and differences of ML models selected
5. Programming
   1. Programming paradigms used
6. Conclusion
7. References
8. Appendix

4. ML

4.1. Project management framework CRISP-DM, KDD and SEMMA.

Which project management framework is best? Regarding CRISP-DM (Cross Industry Standard Process project), KDD (Knowledge Discovery in Databases) and SEMMA (Sample, Explore, Modify, Model, Assess) models, the answer really depends on the nature of the data science project we are working on.

**FIND EXAMPLES AND REFERENCE THEM!!!!**