PROJECT 2: PREDICTING PH

DATA 624 - Predictive Analytics

Group 2

Group Members:

Juliann McEachern

10 December 2019

Contents

Introduction

This project is designed to evaluate production data from a beverage manufacturing company. Our assignment is to predict PH, a Key Performance Indicator (KPI), with a high degree of accuracy through predictive modeling. After thorough examination, we approached this task by splitting the provided data into training and test sets. We evaluated several models on this split and found that **what-ever-worked-best** method yielded the best results.

Each group member worked individually to create their own solution. We built our final submission by collaboratively evaluating and combining each others' approaches. Our introduction should further outline individual responsibilities. For example, **so-and-so** was responsible for **xyz task**.

For replication and grading purposes, we made our code avaliable in the appendix section. This code, along with the provided data, score-set results, and individual contributions, can also be accessed through our group github repository:

- Pretend I'm a working link to R Source Code
- Pretend I'm a working link to Provided Data
- Pretend I'm a working link to Excel Results
- Pretend I'm a working link to Individual Work

1 Data Exploration

We were provided with a file containing 2,571 rows/cases of data and 33 columns / variables.

Response Variable

Understanding the influence PH has on our predictors is key to building an accurate predictive model. PH is a measure of acidity/alkalinity that must conform in a critical range.

Predictor Variables

Text text text.

Data Transformations

Text text text.

2 Predictive Modeling

Text text.

Train

Train text.

Test

Test text.

3 Discussion

Eval text. The end.

4 Conclusion

sfasdfs

Appendix

Code & stuff here.