Bonertz, Brian

DSC550

Week 12: Term Project Final Submission

**SHELF LIFE OF RAW PORK PRIMALS**

**Introduction**

The United States is a global leader in food safety and innovation. Bringing food from the farm to the table in a food safe and efficient way underlies the core values of the food industry. There are countless interventions, processing aids, and critical control points continuously developed to ensure the product delivered to the consumer remains nutritious and wholesome till the point of consumption. While tremendous gains in food safety and quality have been witnessed throughout time with the induction of food safety regulations, sanitation procedures, and chilling and storage practices, challenges remain and efforts toward continuous improvement persist. Due to the very nature of many perishable food items, the industry is in a continuous battle to maintain sufficient quality and wholesome attributes until the point of use.

The meat industry has a particular interest in this challenge. At the moment, of harvest and processing, the product begins the decomposition process. At which point, the purpose to maximize the period of time between packaging and its ends use while maintaining product acceptability becomes a critical economic factor for the processors and retailers. There is great loss, expense, and waste in the disposal of product that is not purchased and consumed before the shelf life expires.

Shelf life is the amount of time that passes before meat becomes unpalatable or unfit for human consumption because of the growth of spoilage organisms. Although the growth of these spoilage organisms is not necessarily considered unsafe, they may present the product organoleptically undesirable to the customer. Organoleptic properties may include overall appearance, purge, surface texture, color, and odor. The purpose of my study is to use these organoleptic properties to evaluate the age and total plate count (TPC) of micros the result in acceptable or unacceptable product condition.

**Scope of the Project**

The data collected is shelf-life data of pork product throughout the 2023 calendar year. With this data, we can evaluate the total micro load of different pork primals at different ages of storage. This allows us to assess the initial load of food spoilage micro-organisms and the generation rate throughout the chilling, and from this data, we can determine the age at which the product begins to become undesirable.

The data used contains and overall assessment of product condition. This product condition contains two variables (acceptable or not acceptable). In this study, I will aim to pinpoint the TPC value and Age at which the product goes from Acceptable product condition to Not acceptable. This study will be assessed on two products. One bone in product type (Back-ribs) and one boneless product type (boneless loins). These particular primals were selected because they are both highly sought after cuts of meat and hold high value in comparison to other cuts of pork. Therefore, establishing conclusion on these two products would supplement value in examining other product types. Providing statistical support in finding the TPC value and Age at which these two products become undesirable would trigger investigation into environmental factors during the production of these products and encourage actions to find process improvement that may extend the shelf life of these products. This would help gain buy-in from key stakeholders to proceed.

**Organized and detailed summary of Milestones 1-3**

Data Preparation Steps:

Step 1: We begin with the initial load of the data. This data has been loaded after the removal of unnecessary columns such as record number, row number, and date / time the record was generated are unique identifiers or have no value to the data we are exploring.

A screenshot of a computer

Description automatically generated

Step 2: Then we change the values of the “Product Condition” column from Acceptable / Not Acceptable to Good / Bad values to better streamline the data and because we already have a column for product age which is the difference between Production\_Date and Sample\_date, these columns can also be removed.

Step 3: We then filter the data down to only include Bone In Back-rib Data and Boneless Loin Data. The below data set is it looks at this point.

A screenshot of a computer

Description automatically generated

Step 4: We want to see how well the organoleptic properties of Odor, Color, and Surface correlate to the overall product condition. The assumption is that product with poor organoleptic properties would result in bad product condition and vise verse. Therefore we separate out these properties, then transform these results into numerical values, and perform a confusion matrix.

A screenshot of a color chart

Description automatically generated

As suspected, we see that good odor, good surface, and good color closely correlate with good product condition. Therefore, we can drop these attributes and focus on Age and TPC values in relation to the overall product condition. The final dataset, after preparation is represented below.

A screenshot of a computer

Description automatically generated

Modeling and Evaluation:

A dataset for each product type was then created. One dataset for the Boneless Loin and one dataset for the Bone In Back-rib so the each product could be evaluated individually. The data was then converted to numerical values and converted into and 80 / 20 split into training and testing data.

Next Logistic regression, KNN classifier, and Random Forest classifier were compared after using a standard scaler in a grid search to find the best fit model for the data. Results revealed that Logistic regression found the best fit. Therefore, Standard Scaler and Logistic regression was used to fit each model and tested to view the accuracy of the model.

Through statistical examination, the accuracy, precision, recall, and f1 score were calculated and displayed below.

Bnls Loins

A number and numbers on a white background

Description automatically generated

BI Backrib

A number on a white background

Description automatically generated

Models were then fitted to S\_Curve charts to get a visual of Age and TPC Values and the rate they transition from acceptable product condition to unacceptable product condition. Each chart contains a horizontal constant line at 0.8 which is highlighting the variable at an 80% confidence level. Therefore, the age or TPC Value the S\_Curve intersects the 0.8 constant line would be our threshold.

Below we see the results for the boneless loin product type.

A graph with blue dots and a red line

Description automatically generated A graph with a red line

Description automatically generated

We can see from the data provided that the S\_curve value does not quite reach the Not Acceptable or Bad product condition with an 80 % confidence level. And the TPC value appears to be just over 8 log as the product reaches an unacceptable level at an 80% confidence level.

Below is the results for the Bone In Backrib product type.

A graph with a red line

Description automatically generated A graph with a red line

Description automatically generated

Here we see a more typical S\_Curve for the Age where the product would be predicted to reach an unacceptable product condition at around 40 days and TPC Value above 8 log with an 80% confidence level.

**Conclusion**

What does the analysis / model building tell you?

We have done a couple of comparisons. We looked to bone in back-ribs verses boneless loin product and examined the relationship of age and TPC value on overall product condition. Results of the analysis indicates that boneless loins have a better shelf life than the bone in back-ribs. The age at which the boneless loin product turns to bad product condition appears to be over 20 days longer than the bone in back-rib product at an 80% confidence level. However, it is interesting that the TPC value of both bone in and boneless product had similar curves. Each indicate a bad product condition around 8.5 log at an 80% confidence level. This would imply that TPC value increases faster on the bone in back-rib product.

Is the model ready to be deployed and what are your recommendations?

I would say that this model is not ready to be deployed for a couple of reasons. First, the accuracy score is a concern to me. The accuracy score of 1.0 can be a result of overfitting. Where the model is not necessarily training from the data but more copying or duplicating the data. So, it may be a benefit to try other models. The other reason would be that further analysis should be done. We have some promising results with the bone in and boneless data on these two product types, it would be good to see how this compares by plant location and other primal types to verify results before it is deployed.

What are some of the potential challenges or additional opportunities that still need to be explored?

Some of the challenges or opportunities to be explored are consistencies in data and data collection. Variables that cause fluctuation in TPC values and organoleptic properties can be difficult to manage. Maintaining proper temperatures and environmental controls are essential to predicting shelf life. In addition, organoleptic evaluation can be very subjective. Product condition of one person may be different than the product condition of another person. Different people may perceive organoleptic properties differently than others. So developing consistency is always a difficult challenge.