**Supplier Pricing Prediction : Outline (rough draft)**

**Scope:**

Caterpillar (construction equipment manufacturer) relies on a variety of suppliers to manufacture tube assemblies for their equipment. These assemblies are required in their equipment to lift, load and transport heavy construction loads. We are provided with detailed tube, component, and annual volume datasets. Our goal is to predict how much a supplier will quote for a given tube assembly.

Each supplier has their unqiue pricing model. Tubes can vary across a number of dimensions, including base materials, number of bends, bend radius, bolt patterns, and end types.

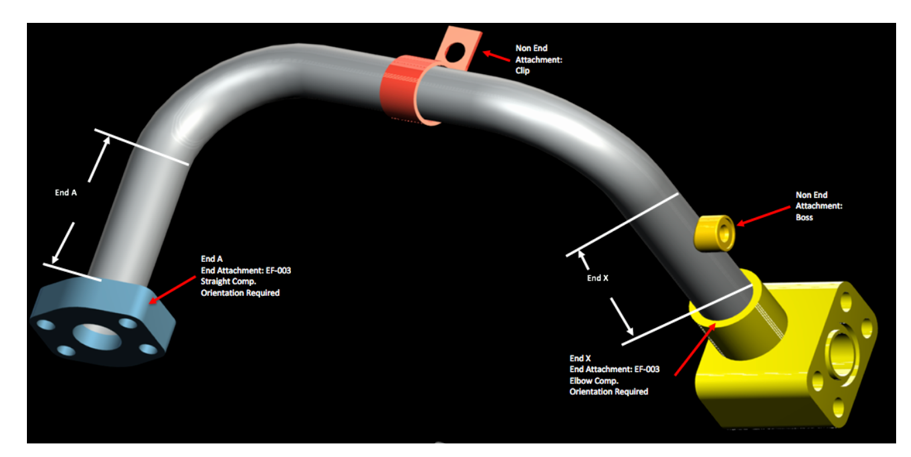
Model Evaluation is to be based on RMSE error. Lowest the better (Predicted -Actual)

**DataSet:**

**Description:**

* Combine characterstics of each tube assembly with Supplier Pricing Dynamics to forecast Quote Price
* Price can be quoted in 2 ways
  + Bracket pricing: Purchase based on Quantity of tubes purchased
  + Non-Bracket Pricing: Purcahse based on minimum\_order Quantity.
* Each Quote has annual usage (how many tube assemblies will be purchased in a given year)

**What does Dataset Contains?**

****

**Dataset Schema: Possible column joins based on common ID’s and component’s description.**

|  |  |  |
| --- | --- | --- |
| **Dataset Tables** | **Content Headers** | **Description** |
| **Tube.csv** | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **tube\_assembly\_id** | material\_id | diameter | wall | length | num\_bends | bend\_radius | end\_a\_1x | end\_a\_2x | end\_x\_1x | end\_x\_2x | end\_a | end\_x | num\_boss | num\_bracket | other | | -Tube Assemblies  -Made of multiple parts  -end\_a\_1x and end\_a\_2x means if end is 1times or 2 times less trhan tube diameter.  -end\_ a and end\_x means type of end connection  -Bosses, brackets, other cutom features that can be attached permanently |
| **Bill of Material** | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **tube\_assembly\_id** | | component\_id\_1 | | quantity\_1 | | component\_id\_2 | | quantity\_2 | | component\_id\_3 | | quantity\_3 | component\_id\_4 | quantity\_4 | component\_id\_5 | | quantity\_5 | component\_id\_6 | | quantity\_6 | | component\_id\_7 | | quantity\_7 | | component\_id\_8 | | quantity\_8 | | | -List of components  - Their Quantities used on each tube assembly |
| **Specs.csv** | |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **tube\_assembly\_id** | spec1 | spec2 | spec3 | spec4 | spec5 | spec6 | spec7 | spec8 | spec9 | spec10 | | **-list of unique specifications for each tube assembly**  **-Used for reference to materials, processes and rust protection etc.** |
| **Tube end form.csv** | |  |  | | --- | --- | | **end\_form\_id** | forming | | **-end types physically formed only using wall thickness** |
| **Components.csv** | |  |  |  | | --- | --- | --- | | **component\_id** | name | component\_type\_id | | **-list of all components used**  **-component category** |
| **Comp\_[type].csv** | **Adopter**   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **component\_id** | | component\_type\_id | | | adaptor\_angle | | overall\_length | | | end\_form\_id\_1 | | connection\_type\_id\_1 | | length\_1 | thread\_size\_1 | thread\_pitch\_1 | | nominal\_size\_1 | | end\_form\_id\_2 | | connection\_type\_id\_2 | | | | length\_2 | thread\_size\_2 | | thread\_pitch\_2 | | nominal\_size\_2 | | | hex\_size | unique\_feature | | orientation | | weight |   **Boss**   |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **component\_id** | component\_type\_id | | type | connection\_type\_id | | outside\_shape | | base\_type | height\_over\_tube | bolt\_pattern\_long | bolt\_pattern\_wide | groove | | base\_diameter | shoulder\_diameter | unique\_feature | | | orientation | weight |   **Elbow**   |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **component\_id** | component\_type\_id | | bolt\_pattern\_long | | bolt\_pattern\_wide | | extension\_length | | overall\_length | thickness | drop\_length | elbow\_angle | mj\_class\_code | | mj\_plug\_class\_code | | plug\_diameter | groove | unique\_feature | | orientation | weight |   **Float**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **component\_id** | component\_type\_id | bolt\_pattern\_long | bolt\_pattern\_wide | thickness | orientation | weight |   **HFL**   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **component\_id** | component\_type\_id | hose\_diameter | corresponding\_shell | coupling\_class | material | plating | orientation | weight |   **Nut**   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **component\_id** | component\_type\_id | hex\_nut\_size | seat\_angle | length | thread\_size | thread\_pitch | diameter | blind\_hole | orientation | weight |   **Other**   |  |  |  | | --- | --- | --- | | **component\_id** | part\_name | weight |   **Sleeve**   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **component\_id** | component\_type\_id | connection\_type\_id | length | intended\_nut\_thread | intended\_nut\_pitch | unique\_feature | plating | orientation | weight |   **Straight**   |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **component\_id** | component\_type\_id | bolt\_pattern\_long | bolt\_pattern\_wide | head\_diameter | overall\_length | thickness | mj\_class\_code | groove | unique\_feature | orientation | weight |   **Tee**   |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **component\_id** | component\_type\_id | bolt\_pattern\_long | bolt\_pattern\_wide | extension\_length | overall\_length | thickness | drop\_length | mj\_class\_code | mj\_plug\_class\_code | groove | unique\_feature | orientation | weight |   **Threaded**   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **component\_id** | | component\_type\_id | | | | adaptor\_angle | | overall\_length | | | hex\_size | | | end\_form\_id\_1 | | connection\_type\_id\_1 | | | | length\_1 | thread\_size\_1 | | thread\_pitch\_1 | | nominal\_size\_1 | | end\_form\_id\_2 | | connection\_type\_id\_2 | | | length\_2 | thread\_size\_2 | | | thread\_pitch\_2 | | | nominal\_size\_2 | | | end\_form\_id\_3 | | | connection\_type\_id\_3 | | | | | length\_3 | thread\_size\_3 | | | thread\_pitch\_3 | | nominal\_size\_3 | | end\_form\_id\_4 | | connection\_type\_id\_4 | | length\_4 | | thread\_size\_4 | | | thread\_pitch\_4 | | nominal\_size\_4 | | | | unique\_feature | | | orientation | | | weight | | | **-information about each components** |
| **Type\_[type].csv** | **Component**   |  |  | | --- | --- | | **component\_type\_id** | name |   **Connection**   |  |  | | --- | --- | | **connection\_type\_id** | name |   **End**   |  |  | | --- | --- | | **end\_form\_id** | name | | **-contains name sfor each feature** |
| **Train.csv** | |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | **tube\_assembly\_id** | supplier | quote\_date | annual\_usage | min\_order\_quantity | bracket\_pricing | quantity | cost | | **-train rows: 30,213**  **-total rows: 60,448**  **Train Split: 49.9%** |
| **Test.csv** | |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | **id** | tube\_assembly\_id | supplier | quote\_date | annual\_usage | min\_order\_quantity | bracket\_pricing | quantity | | **-test rows : 30,235**  **Total wors: 60,448**  **Test Split: 50%** |
| **Submission Sample.csv** | |  |  | | --- | --- | | **id** | cost | |  |

**Primary Join: Tube Assembly ID**

**Secondary Join: Component ID**

**Join Summary Table:**

|  |  |
| --- | --- |
| **Datset tables** | **Common Keys** |
| Connect Tube with Specs | Use Material ID with Specs. This has common material ID’d. |
| Connect Tube with bill of material | Use common Tube assembly ID |
| Connect Bill of material with components, component types, types of (component, connection and end), and tube end form | Use Common ID’d (Components ID, Component Type ID, end form ID, connection type ID) |
| Connect combined features with train table.  Train and Evlaute Table | Use common tube assembly ID |
| Use trained model to predict cost using Test dataset. | Use common tube Assembly ID |

**Challenge: How to connect all 21 tables in a way that we can draw insights?**

**Example:** Tube Assembly ID (Primary Join) is not directly linked to Component ID (Secondary join).

It is connected to multiple Component ID’s (1 to 8) in BOM table, which has to be linked to 11 other component tables.

**Why linking tables important?**

* Need to identify how selection of features such as use of certain material type, number of components in the tube assembly, volume etc will predict supplier pricing.
* Lower the difference between predicted pricing and given pricing in terms of RMSE, better should be the model predictability.

**Other Ideas:**

* Use external web sources such as material price index to understand pricing trends
* Segment suppliers based on order date, (volume or frequency) and predicted cost.
* Compare prediction with deep learning model.