Technology used: Python

Assumptions:

1. Assumed Expected Revenue and Expected Conversion are accurate.
2. No bid means 0 $. (converted Nan / None values to zero in bid Price).
3. Converted Bid price into category values from float for doing modelling.

Steps:

1. Checked the volume of leads if they are in range or not. (They are in range, aliter way of doing it is calculated volume +5%, volume-5%.
2. Total Revenue for each bid is ExpectedConversion \* ExpectedRevenue ( as number of leads and loan = 1 for each bid).
3. Net Revenue = Total Revenue (for each bid) – BidPrice.
4. Sanity check of Total Net Revenue for Bids that are accepted is calculated and is found to be 1207654.0432900002.
5. Contribution of each bid price towards NetRevenue is visualized in Pie Chart. (75 $ bid price contributes 55.63 % towards total Net Revenue where as 3 $ bid price contributes only 2 %.)
6. Simple model as explained in the document is implemented and check with bid Prices. The accuracy using Simple model is 100 %.
7. Tested with a bunch of machine learning models like LR, LASSO, EN, KNN, Decision Trees and SVM using regression with negative mean squared error as metric.

A box and whisker plots are plotted to compare the ML algorithms. KNN (Nearest neighbours algorithm performs best among them ).

Thanks