**Manual Features Creation:**

**Text Features:**

* Extreme positive words:
  1. Extract the extreme positive words from each review
  2. Rank the words:
     + 3- perfect, amazing, awesome, excellent, outstanding, beyond fabulous, incredible, superb, top class, stunning, Wow
     + 2- really, very, great, Everything
     + 1- Good, helpful, friendly, spacious, welcoming
     + 0- no extreme words
  3. Create 3 binary features – each review get 1 if used at least one words from the above list for each group of words
* Extreme negative words:
  1. Extract the extreme positive words from each review
  2. Rank the words:
     + 3- worst, Shockingly, extremely, terrible, disgusting, incredibly, Completely, Surprised, disappointed, everything, exceptional
     + 2- too, poor, really, very
     + 1- small, not good, basic, not great, expensive, could be better, claustrophobic
     + 0- no extreme words
  3. Create 3 binary features – each review get 1 if used at least one words from the above list for each group of words
* Positive and negative length:
  1. Find the length of the positive and negative parts of the review
  2. Rank the lengths:
     + 0- less than 100 characters
     + 1- 100-199 characters
     + 2- more than 200 characters
* Positive and negative length proportion:
  1. Find the length of the positive and negative parts of the review and calculate the proportion: positive\_len/negative\_len
  2. Rank the lengths:
     + 0- <0.7
     + 1- 0.7-4
     + 2- >4
* Previous round lottery result:
  1. low: x<3
  2. med1: 3<=x<=5
  3. med2: 5<x<8
  4. high: x>=8
* seconds on page (DM page):
  1. low: x<15
  2. med: 15<=x<30
  3. high: x>=30

**Global Features:**

* Gender
* History\_decisions: The proportion of rounds the DM chose the hotel
* history\_lottery\_result\_high: The proportion of rounds the results of the lottery was higher than 8 🡪 the payoff was positive
* history\_lottery\_result: the average lottery results
* history\_chose\_lose: The proportion of rounds the DM chose the hotel and lose (result<8)
* history\_chose\_earn: The proportion of rounds the DM chose the hotel and earn (result>=8)
* history\_not\_chose\_lose: The proportion of rounds the DM chose the hotel and could lose (result<8)
* history\_10\_result: The proportion of rounds the lottery results was 10

\*\* proportion might the average of the weighted average with where i=number of rounds between the current round the data’s round

**Previous round vs all history vs no history:**

* Previous round: use the text, decision and feedback from one previous round.
* All history (NN): use the text, decision and feedback from all the previous rounds.
* All history text (no NN): use the calculated history text features
* All history (no NN): use the calculated history not text features (global features from above)
* No history: use only the text from the current round to predict the current decision
* No text: use only the numeric features