316 Group Project

- requirement:
  - As part of the group decisions, each student will be allocated to one of the following tasks:
    - Descriptive analysis of the dataset + Error analysis
    - Preprocessing + Literature review
    - Implementation + Results
- Analysis
  - Data Collection and Preparation
    - the sentiment dictionary
      - Lexicon
        - Negation rule
        - Capitalisation rule
        - Intensifier rule
        - Diminisher rule
        - Exclamation rule
        - 😐？？ (keep or delete)
      - Intensifier
      - Exclamation
  - Feature Extraction and Representation
    - **cross-validation:**

    - Model Selection and Training
        - Bayes,
            - naive Bayes ： likelihood
                - should we need apply a **smoothing** technique

$$p(t_j|s_i) = \frac{count(t_j, s_i) + 1}{count(s_i) + |V|}$$

        - support vector machine (SVM),
        - logistic regression,
        - deep neural network,
    - Sorting and categorising through fractions
    - Type Evaluation and Tuning
- key question:
    - the sentiment dictionary
    - sorting（Dealing with adjectives and adverbs of degree）
    - neural network or the model choose
- Relevant methods
    - Bing Liu’s model