- Automated essay grading
    - The purpose of this project is to implement and train machine learning algorithms to automatically assess and grade essay responses.
    - Dataset: Essays with human graded scores
- Sentence to Sentence semantic similarity
    - Can you identify question pairs that have the same intent or meaning?
    - Dataset: Quora question pairs with similar questions marked
- Fight online abuse
    - Can you confidently and accurately tell whether a particular comment is abusive?
    - Dataset: Toxic comments on Kaggle
- Open Domain question answering
    - Can you build a bot which answers questions according to the student's age or her curriculum?
    - Facebook's FAIR is built in a similar way for Wikipedia.
    - Dataset: NCERT books for K-12/school students in India, NarrativeQA by Google DeepMind and SQuAD by Stanford
- Social Chat/Conversational Bots
    - Can you build a bot which talks to you just like people talk on social networking sites?
    - Reference: Chat-bot architechture
    - Dataset: Reddit Dataset
- Automatic text summarization
    - Can you create a summary with the major points of the original document?
    - Abstractive (write your own summary) and Extractive (select pieces of text from original) are two popular approaches
    - Dataset: CNN and DailyMail News Pieces by Google DeepMind