a) NLP application used for allocating folders for mails.

In NLP, the emails are automatically parsed and analyzed using text classification and keyword extraction methods.

Text classification is of two types: - rule-based and machine learning based text classification. Rule techniques use a set of manually constructed language rules to categorize text into categories or groups. The rule will classify the content of the text by using semantic text elements.

In machine learning it uses training and prediction phases. First the raw data is preprocessed. The different steps include tokenization, text-cleaning, POS tagging, stopwords and lemmetization.

b) Awarding grade to the essay question

Automated essay scoring process is used to award the grade. This evaluation is done by computers where grading models are learned using essay datasets scored by different graders. Other features used to evaluate the essays are lexical diversity, sentence count, word frequency, word count, average length, structure and organization of an essay. Most of the algorithms are regression based where the essay scores as a range of values and predict a floating point value within that range.

c) Model used by doctors to predict the diagnosis.

Fuzzy Reasoning module is used to read and interpret the responses from the user, track and monitor all the symptoms that the user has already responded to and to administer questions to the user that are most relevant based on the dataset of diseases that is maintained. Each disease is modeled as a bucket, where each bucket is associated with a symptom. The fuzzy SVM classifier was trained using a set of training documents and then converted into feature vectors. The feature extracted from the user answers that are passed to the fuzzy SVM model suggests the diagnosis by performing the classification on those words contained in the patient answer.