Compulsory Task 1

Now that you know a little about the various categories of NLP applications, do you think you can identify these when considering particular applications? Give it a try.
In a file called nlp_1.pdf, categorise each of the following use-cases:
(Use the categories we have discussed on this boot camp so far)

a. A model that allocates which mail folder an email should be sent to (work, friends, promotions, important), like Gmail's inbox tabs.

- Text classification - sorting emails into categories.
- This has also been used as an example for classification when discussing classification vs. regression.
- Semi-supervised learning algorithm: it should detect patterns in unlabelled data (unsupervised), but can use previously labelled data to classify those patterns (supervised).

b. A model that helps decide what grade to award to an essay question. This can be used by a university professor who grades a lot of classes or essay competitions.

- Automatic summarisation: If the answers are very long, a summarisation can make the job easier.
- Regression: Give a numeric score to the essay depending on the amount of correct information in the essay. (Scores 0-100)
- OR Classification: Give a score in the A-F way. Each letter represents a category with A being the best and F being the worst (E does not exist).
- Semi-supervised learning algorithm: The professor can use previously graded essays as the labelled data to train the algorithm. This can then help the algorithm learn what length, structure, arguments etc to look for (supervised). It can use this data to find patterns in new essays and find different patterns as well using an unsupervised algorithm.

c. A model that provides assistive technology for doctors to provide their diagnosis. Remember, doctors ask questions, so the model will use the patients' answers to provide probable diagnoses for the doctor to weigh and make decisions on.

- Text Classification - to use the information provided by patients to classify into different symptoms and diagnoses
- Question Answering - ask question: 'which disease' and receive an answer
- Semi-supervised learning algorithm - algorithm has to learn from humans and previous datasets: patients answer questions about health, symptoms get filtered out of those answers, diagnoses are given (in a classification style). This is the supervised learning part.

This information can be used to find further patterns etc in the new patients' answers (unsupervised).

Source: The questions were answered using the material we have been previously given: Machine Learning PDF by Hyperion DEV and NLP Applications by Hyperion DEV.