Reading Tasks ▪ The Scikit-learn library has a module ([sklearn.feature\_selection](https://scikit-learn.org/stable/modules/feature_selection.html)) for feature selection.

Study all the feature selection methods in the module; at the end of lectures for this module, you will be given a dataset whose dimensionality you will be required to reduce using the library. ▪ Read and make a brief presentation about Feature Extraction Techniques (Principal Component Analysis and Linear Discriminant Analysis)

▪ Field attachment is defined, in Makerere University, as the field based practical work carried out by staff and students for the purpose of teaching and/or research in places outside of the University control but where the University is responsible for the necessary safety of its staff, students, and others exposed to their activities.

▪ During Field Attachment, students get attached to organizations to experience the real life of work. Overtime, Makerere University has placed its interns Field Supervisors in organizations of attachment. These Field Supervisors have provided valuable feedback to the University about its students. Unfortunately, this feedback is in form of unstructured text and large in volume which makes it difficult to process and gain useful insights.

▪ You are required to download this dataset from https://www.fams-cit.com/fscomments and perform the following text analysis tasks of the dataset

▪ Create a corpus from the downloaded comments from Field Supervisors.

▪ Cluster the comments of Field Supervisors about Interns into categories: Excellent, Good, Neutral, Poor, Very Poor.

▪ Assess the performance of your clustering algorithm above.

▪ Create a Named Entity Recognition (NER) model that takes in a comment as an input and outputs the Entities, if any, belonging to the categories: Person, Organization, Place/Location, Time

▪ Create a visualization to show insights about the dataset. (Create the visualization using D3 and host it at the same server you have used in the previous assignments)