Nitte Meenakshi Institute of Technology
Department of Computer Science and Engineering
18CS54 Data Mining
Project Guidelines and Templates

Due Dec 29,2021 Wednesday @ 11:55 pm – Project Proposal
Jan 17, 2022, Monday @ 11:55 pm – Project Report and supporting artifacts.

Overview

The goals of this learning activity 2 are

- Apply data mining techniques learnt for a chosen real-world problem and dataset (*Refer sources like: Kaggle, kdnuggets, UCI etc*).
- As a team of 2-4 members, you can discuss your findings on the chosen questions and consolidate your learnings as a report (Answers to each question: Algorithm, Working screen shots and results and Analysis).
- Make a consolidated report about your findings
- This learning activity is an important part of the course and counts for 10% of your final grade. Grades will be based on the problem chosen, data mining approaches used, completeness of your findings, analysis, quality of the proposal and report.

Grading

- This learning activity is worth for 10 marks
- Marking Scheme with timelines

Document	Weightage	Due Date	Comments
Project Proposal		29-Dec-2021	Refer proposal
Project Report	10	17-Jan-2021	Refer report

- Late delivery without prior notification and permission from the course coordinator will result in a loss of 10% of the marks per day.

GitHub collaboration for submission

- Identify the team members and team leader. Team leader can create GitHub account and add rest of the members including course coordinator as a collaborator to access/update the project. Team details and GitHub invite must be sent to the course coordinator on or before the submission deadlines.
- All the upload and submissions can be made through the GitHub.

Course Coordinator: Dr.Vani V

<div align="center">

Project Proposal
<Title of Project>
<Member IDs>, <Member Names>

</div>

Introduction

Provide a brief overview of data mining. Describe what your proposal is about and the organization of the rest of the proposal. Include whether you will be performing data mining tasks, implementing a new algorithm in Weka (or another data mining tool), or modifying some other system to incorporate data mining features, etc. Basically, provide the nature of your project.

Data Mining Task

Provide the specific tasks you will perform on the data set. **Include specific questions you will investigate, and the goals for the tasks.** This should be independent of the specific techniques you will use to achieve your goals.

Data Set

Describe the data set(s) you will be using in your project. Include the origin of the data set, an overview of the data set organization, attributes of the data, and challenges of the data set you've selected. Include any information you have about missing values in the data set.

Methods and Models

Describe in detail the data mining methods and models you plan to employ to achieve the goals you set in the *Data Mining Task* section of your document. Include some mention of necessary data transformation. If you're implementing a technique, you should have some idea of how it will be implemented and incorporated in the chosen language (Python/R). If you are combining techniques, explain how you intend to use the output of one technique as input into another technique.

Assessment

Discuss the assessment methodology you will use to validate that you have found meaningful patterns.

Presentation and Visualization

Describe how your results will be presented and visualized in such a way to show meaningful patterns in the data.

Roles

In this section, discuss the roles that each group member will have in the project.

Schedule

The schedule is a table of dates and tasks that you plan to complete by those dates. Tasks to be done by the progress report must be listed, as well as any other dates you want to set for yourselves. Additional deadlines are highly recommended. Be sure to include when you will complete data transformation, modelling, assessment, visualization, etc.

Date	Tasks to be Completed
??/??/21	Tasks completed by chosen date
??/??/22	Tasks to be completed by the final report date
??/??/22	Tasks completed by the class presentation

Bibliography

This is where you list bibliographic information for any references you made throughout the proposal. You should have 5 – 10 references.

Table of Contents

Course Coordinator: Dr.Vani V