

Lecture 13:

Project Discussion and Update

COSC 526: Introduction to Data Mining
Spring 2020



THE UNIVERSITY OF
TENNESSEE
KNOXVILLE
BIG ORANGE. BIG IDEAS.®

Project: Status

Project

Motivation Describe the motivation of your work. To build the motivation, you can answer these questions:

- What is the problem you are tackling?
- How is the problem solved today?

Contributions List between 2 and 4 contributions of your work. Contributions are bullet points that define your solution. E.g.,

- We build a system that
- We validate the system accuracy by
- We measure the performance of the system by ...
- Write a section of 150 - 200 words

Project

Tests List the type of tests (measurements) you will perform. E.g.,

- What are your metrics of success?
- Where do you run your tests?
- What tests do you perform?
- How many times do you run each test?
- What do you measure?
- Write a section of 250 - 350 words.

Milestones

- March 26: **Define your project (DONE)**
 - Describe the motivation of your work
 - List between 2 and 4 contributions of your work
 - List the type of tests (measurements) you will perform
- April 2: **No lecture (DONE)**

Milestones (DONE)

- **April 9: Create a new notebook with your solution**
 - Write down the steps of your solution in distinct text cells; add one or multiple cells (as needed) to hold your code for each step. You can leave these software cells empty for the moment. Expand the text cells describing your solution.
 - Add visualization cells that allow you to visualize results. You can leave these software cells empty for the moment.
 - Add software to the code cells that upload data from source and pre-process data.
 - Push your notebook into your GitHub repository as frequently as needed

Milestones

- April 16: Finalize software and complete the test run within your notebook (DONE)
- April 23: Create your poster and get feedback, submit draft
- April 30: Submit your final notebook and poster in GitHub
- May 7: Submit your 2-page abstract in GitHub and video in YouTube

Project progress

Project Titles

- Levente Dojcsak: Predicting the Development of CKD and Identifying Preventative Treatments
- Konstantinos Georgiou: Analyzing and predicting bottlenecks on the distribution of COVID-19 Vaccines
- Michael Wermert: Stock Predictor: Using Machine Learning to Predict Stock Market Behavior
- Smit Patel: Detecting Malware Using Machine Learning

Project Titles

- Georgia Channing: Traces in the Noise: Identifying Invalid Vessel Paths
- Anuj Gautam: Add a symbol and try again: A comprehensive study of password policies
- Mirka Mandich & Jake Maeker: Machine Learning Applied to HIT-SI Spheromak Data
- Carter White: Impact of Champion Selection on League of Legends Rank

Project Titles

- Xinlan Jia & Candice Chen: US Airbnb Price Prediction
- Zhixiu Lu: Using Microbiomes to Predict Environmental Factors via Machine Learning Approaches
- Azarang Asadi: Motor control quantification using lower-limb body kinematics
- Tommy Wong: Applications of Variational Autoencoders in Analyzing Ferroelectric Domains

Project Titles

- Jerome Kovoor and Shree Neupane: Analyzing the effect of climate change on global food production using K-means clustering or DBSCAN
- Fabian Fallas Moya: How many are good enough?
Finding the best number of annotated images into a self-training algorithm
- Gerald Jones: Subgroups and Factors of ESRD

Milestones

- April 16: **Finalize software and complete the test run within your notebook (DONE)**
- **April 23: Create your poster and get feedback, submit draft**
- **April 30: Submit your final notebook and poster in GitHub**
- **May 7: Submit your 2-page abstract in GitHub and video in YouTube**