Data-Driven Course Insights: Predicting Grade Trends

Project Check In - Team 5

Team 5

Name	Email	GitHub Handle
Arlette Diaz	adiaz218@uic.edu	adiaz218
Marianne Hernandez	mhern85@uic.edu	marhern19
Nandini Jirobe	njiro2@uic.edu	nandinijirobe
Sharadruthi Muppidi	smuppi2@uic.edu	sharadruthi-uic
Sonina Mut	smut3@uic.edu	snina22
Yuting Lu	lyuti@uic.edu	yutinglu103

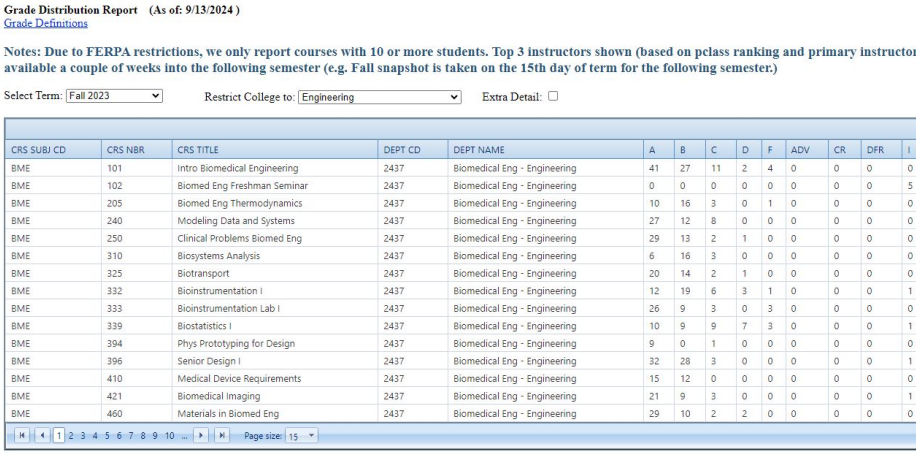
Our Big Idea 

The big idea of this project is to predict course grade distributions and popularity rankings for upcoming semesters, enabling students to make informed decisions about their class selections. By shifting the focus from individual grade predictions to overall course outcomes, the project provides insights into course grading trends and demand. It uses clustering to rank courses based on student performance and popularity, and topic-based grouping to help students discover courses aligned with their interests, factoring in professor expertise and class attributes. This data-driven tool uncovers hidden patterns, aiding both students and academic planning.

Dataset #1: UIC Grade Distribution App

Features: Department, Course Number, Course Title, Department CD, Department Name, Grade Distribution (A, B, C, D, F), number of registered students, semester offered, number of withdrawals, etc.

Cleaning: Removing irrelevant features such as CRN, grouping course numbers to just 100, 200, 300, 400, 500 level classes, converting all distributions to percentages, etc



Grade Distribution Report (As of: 9/13/2024)

Grade Definitions

Notes: Due to FERPA restrictions, we only report courses with 10 or more students. Top 3 instructors shown (based on pclass ranking and primary instructor available a couple of weeks into the following semester (e.g. Fall snapshot is taken on the 15th day of term for the following semester.)

Select Term: Fall 2023 Restrict College to: Engineering Extra Detail:

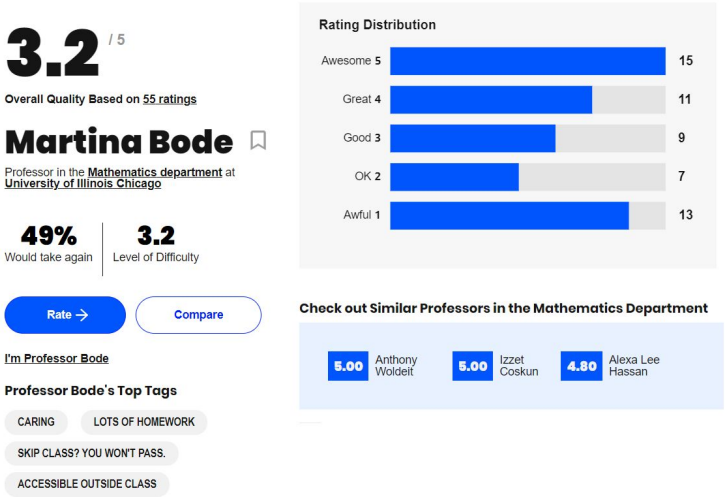
CRS SUBJ CD	CRS NBR	CRS TITLE	DEPT CD	DEPT NAME	A	B	C	D	F	ADV	CR	DFR	I
BME	101	Intro Biomedical Engineering	2437	Biomedical Eng - Engineering	41	27	11	2	4	0	0	0	0
BME	102	Biomed Eng Freshman Seminar	2437	Biomedical Eng - Engineering	0	0	0	0	0	0	0	0	5
BME	205	Biomed Eng Thermodynamics	2437	Biomedical Eng - Engineering	10	16	3	0	1	0	0	0	0
BME	240	Modeling Data and Systems	2437	Biomedical Eng - Engineering	27	12	8	0	0	0	0	0	0
BME	250	Clinical Problems Biomed Eng	2437	Biomedical Eng - Engineering	29	13	2	1	0	0	0	0	0
BME	310	Biosystems Analysis	2437	Biomedical Eng - Engineering	6	16	3	0	0	0	0	0	0
BME	325	Biotransport	2437	Biomedical Eng - Engineering	20	14	2	1	0	0	0	0	0
BME	332	Bioinstrumentation I	2437	Biomedical Eng - Engineering	12	19	6	3	1	0	0	0	1
BME	333	Bioinstrumentation Lab I	2437	Biomedical Eng - Engineering	26	9	3	0	3	0	0	0	0
BME	339	Biostatistics I	2437	Biomedical Eng - Engineering	10	9	9	7	3	0	0	0	1
BME	394	Phys Prototyping for Design	2437	Biomedical Eng - Engineering	9	0	1	0	0	0	0	0	0
BME	396	Senior Design I	2437	Biomedical Eng - Engineering	32	28	3	0	0	0	0	0	1
BME	410	Medical Device Requirements	2437	Biomedical Eng - Engineering	15	12	0	0	0	0	0	0	0
BME	421	Biomedical Imaging	2437	Biomedical Eng - Engineering	21	9	3	0	0	0	0	0	1
BME	460	Materials in Biomed Eng	2437	Biomedical Eng - Engineering	29	10	2	2	0	0	0	0	0

1 2 3 4 5 6 7 8 9 10 ... Page size: 15

Dataset #2: Rate My Professor

Features: Average rating provided by students

Cleaning: We will be scraping the average rating for the professor from RMP. Hence we will need to remove any excess data we collect. For professors who do not have a rating yet, their default rating will be a 5



Dataset #3: Class Scheduler Data

Features: Course titles, Descriptions, Credit hours, Enrollment restrictions, Class types (LAB, LEC), Class Timings, Meeting Days, Class Location, Instructor details, Instructional methods.

Cleaning: Removing irrelevant features such as CRN, room, building, categorizing class times to morning, afternoon, evening, grouping class types to just lab/lecture

CS 109

Programming for Engineers with MatLab

3 hours. Credit is not given for CS 109 if the student has credit for CS 111 or CS 112 or CS 113. Extensive computer use required. Prerequisite(s): Credit or concurrent registration in MATH 180. To be properly registered, students must enroll in one Laboratory-Discussion and one Lecture-Discussion.

CRN	Course Type	Start & End Time	Meeting Days	Room	Building Code	Instructor	Meets Between	Instructional Method
30607	LBD - BAA	08:00 AM - 09:50 AM	R	2249E	2SELE	Riazi, S		Meet on campus
30608	LBD - BAB	10:00 AM - 11:50 AM	R	2249E	2SELE	Riazi, S		Meet on campus
30609	LBD - BAC	12:00 PM - 01:50 PM	R	2249E	2SELE	Riazi, S		Meet on campus
30610	LBD - BAD	02:00 PM - 03:50 PM	R	2249E	2SELE	Riazi, S		Meet on campus
30941	LBD - BAE	04:00 PM - 05:50 PM	R	2249E	2SELE	Riazi, S		Meet on campus
30942	LBD - BAF	08:00 AM - 09:50 AM	F	2249E	2SELE	Riazi, S		Meet on campus
36426	LBD - BAG	10:00 AM - 11:50 AM	F	2249E	2SELE	Riazi, S		Meet on campus
36427	LBD - BAH	12:00 PM - 01:50 PM	F	2249E	2SELE	Riazi, S		Meet on campus

Dataset #4: Google Scholar

Features: Professor name, professor research studies and papers history, professor role

Cleaning: We will need to scrape the data for each professor, calculate the similarity between the professor's research and course they are teaching as a percentage

Sathya N. Ravi

FOLLOW

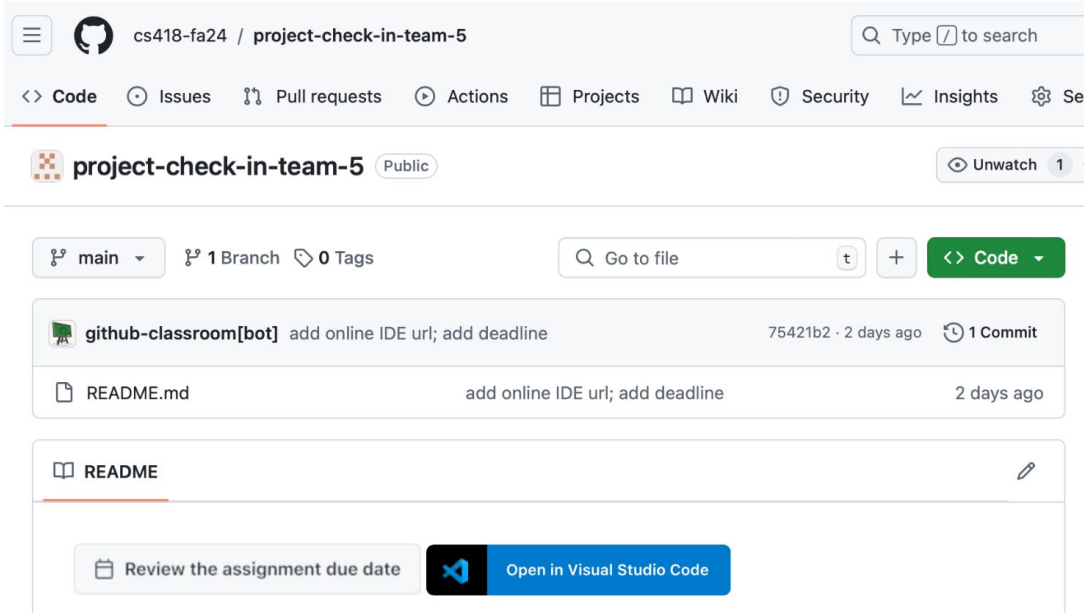
Assistant Professor Of Computer Science at University of Illinois at Chicago
Verified email at uic.edu - Homepage

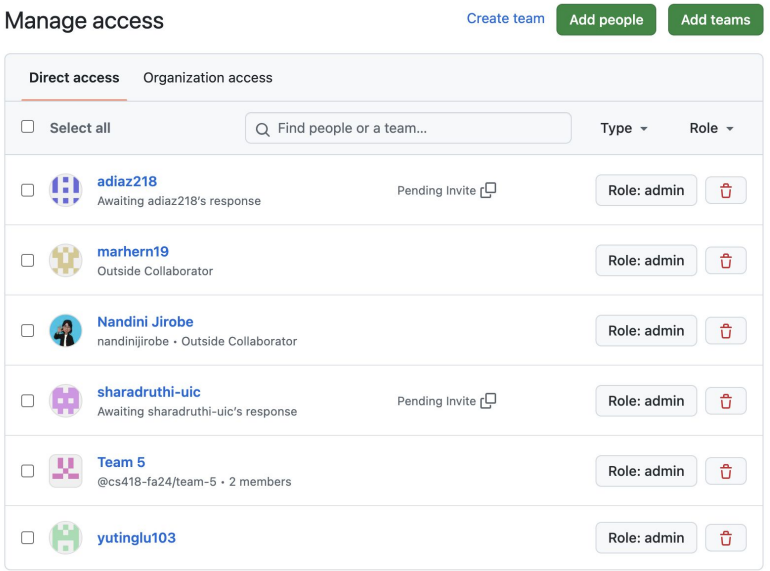
Optimization Machine Learning Computer Vision

TITLE	CITED BY	YEAR
Deep unlearning via randomized conditionally independent hessians R Mehta, S Pal, V Singh, SN Ravi Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern …	75	2022
Tensorize, factorize and regularize: Robust visual relationship learning SJ Hwang, SN Ravi, Z Tao, HJ Kim, MD Collins, V Singh Proceedings of the IEEE Conference on Computer Vision and Pattern …	72	2018
Explicitly imposing constraints in deep networks via conditional gradients gives improved generalization and faster convergence SN Ravi, T Dinh, VS Lokhande, V Singh Proceedings of the AAAI conference on artificial intelligence 33 (01), 4772-4779	71 *	2019
Fuzzy assessment of FMEA for rotary switches: a case study S Vinodh, S Aravindraj, SN Ravi, N Yogeshwaran The TQM Journal 24 (5), 461-475	45	2012

Solution

We will gather data on student enrollment, graduation rates, GPA distributions, retention, and demographic variables. The project will compare traditional in-person learning, online learning, and hybrid models to see how they influence student performance. The project aims to derive insights that can be applied by academic institutions to optimize learning formats based on student demographics and performance data. Part of the scope is to provide guidance on how institutions might allocate resources effectively to improve retention and graduation outcomes. The end result of this project should be a set of clear, data-driven insights that can inform decisions on improving learning experiences for different student demographics, thus enhancing retention and graduation rates.





Github Repository Creation Proof and Link to Repository