# Applying Counterfactual Inference Using Predictive Models to Determine Alternative Course Sequences for Helpable Students

Christopher Kott, Zachary Meyer, Lincoln Schroeder, and Nicholas Tiede

**Faculty Mentors:  Douglas Selent**
University of Wisconsin - Platteville

## Abstract

The purpose of this research is to provide data and tools necessary for students and faculty advisors to determine alternative course sequences for helpable students who are struggling on their current course sequences.

## Research Questions

1) Are there common course sequences between students?
2) Can we determine possible alternative sequences?
3) How accurately can we predict alternate course sequences grades?
4) How many struggling students are helpable?

## Data

Our dataset consists of the historical grade data of 689 unique students from University of Wisconsin - Platteville graduates and withdrawals between the years 2013 - 2018.

**Data Sample**

| Anonymized student ID | Year | Course Name | Grade | Academic Standing | Course Credits |
|---|---|---|---|---|---|
| 1604 | 2017 | Artificial Intelligence | C | Good | 3 |
| 6166 | 2015 | Intermediate Software Engineering | C+ | Good | 3 |
| 3313 | 2013 | Data Communications | D | Probation | 3 |

## Model Performance Experiment Design

- Five-fold stratified (on target letter grades) student-level cross-validation
- Predicting future student course grades
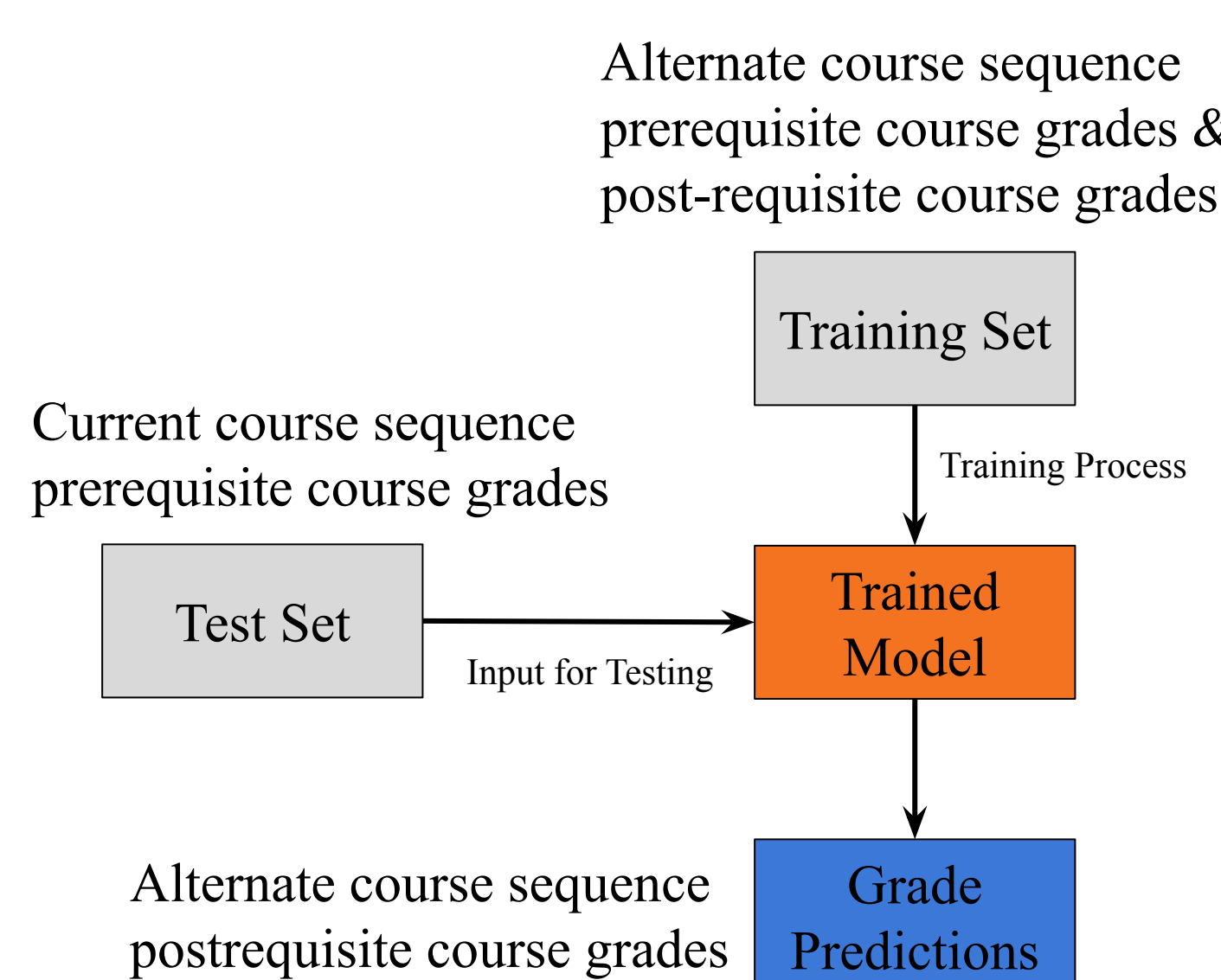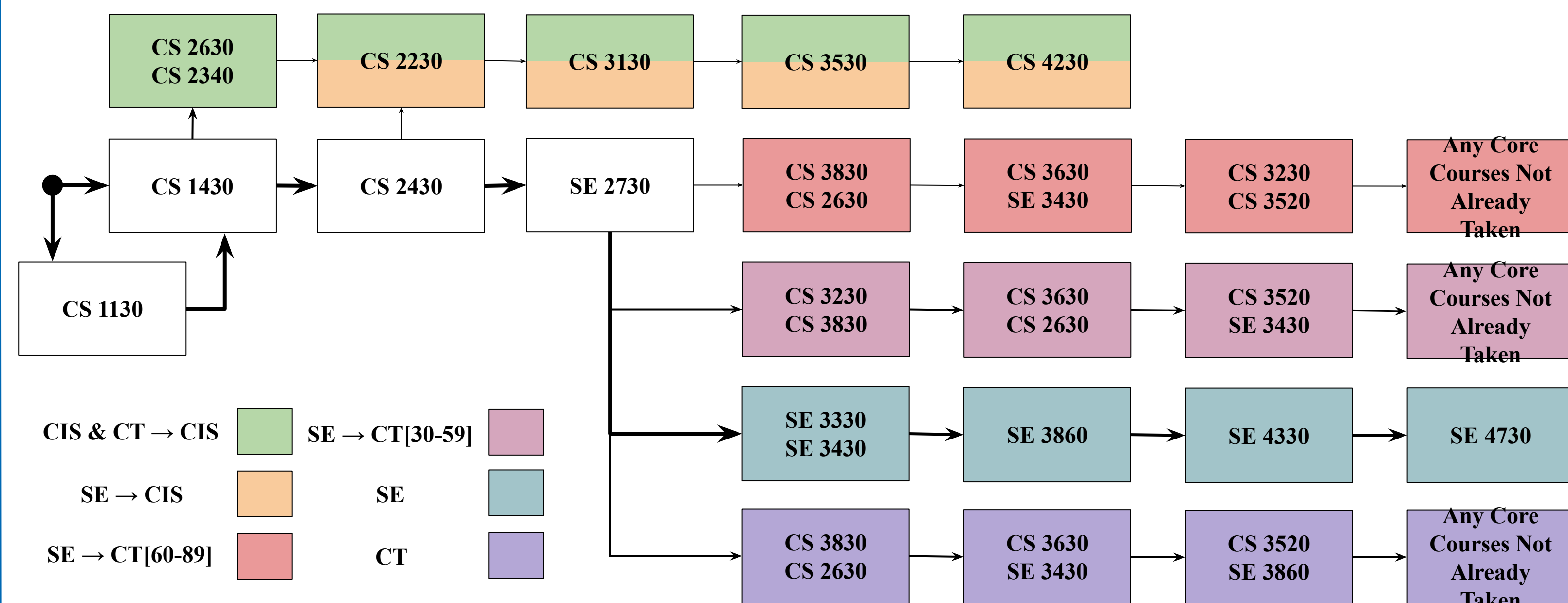- Trained on three different amounts of past data



**All Prerequisites**   **Immediate Prerequisites**   **Root Prerequisites**
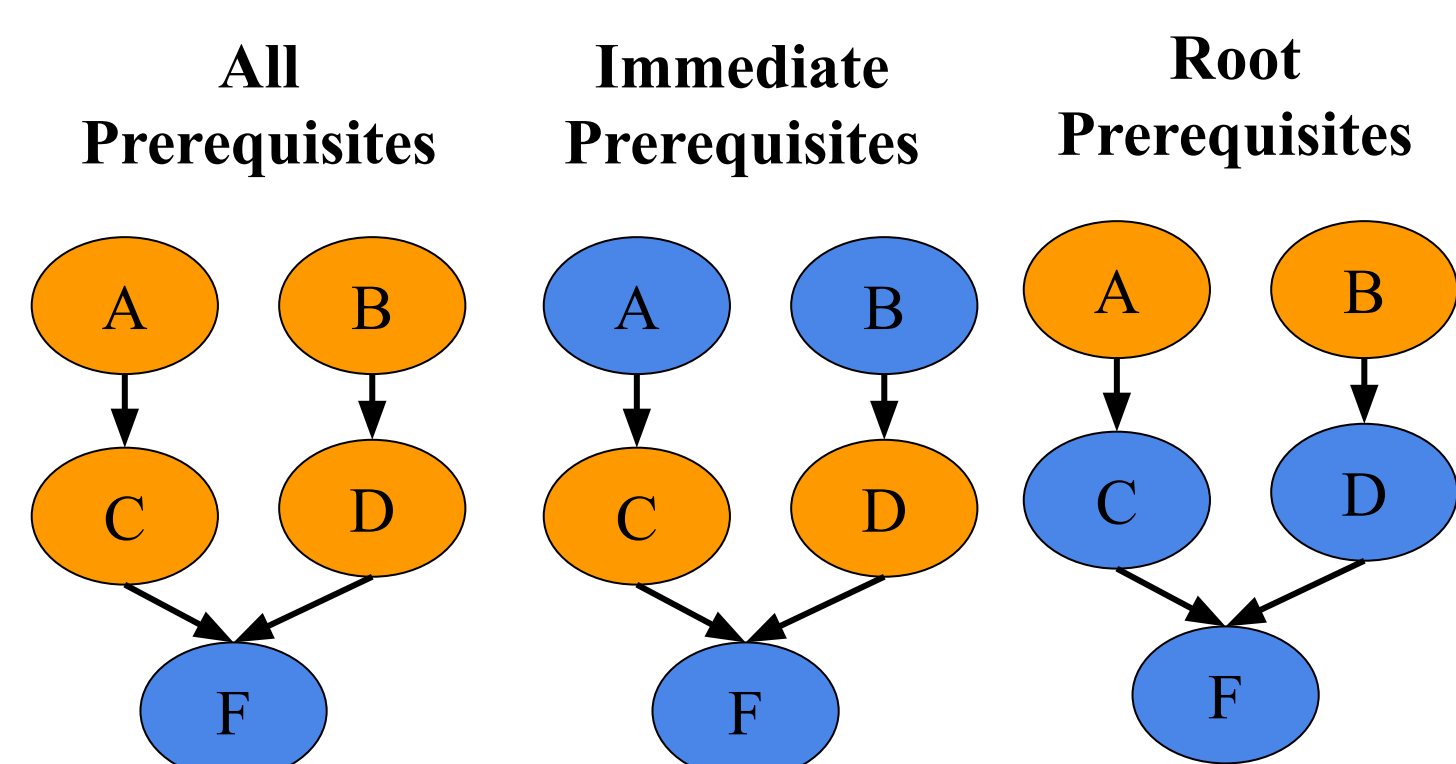
## Counterfactual Inference
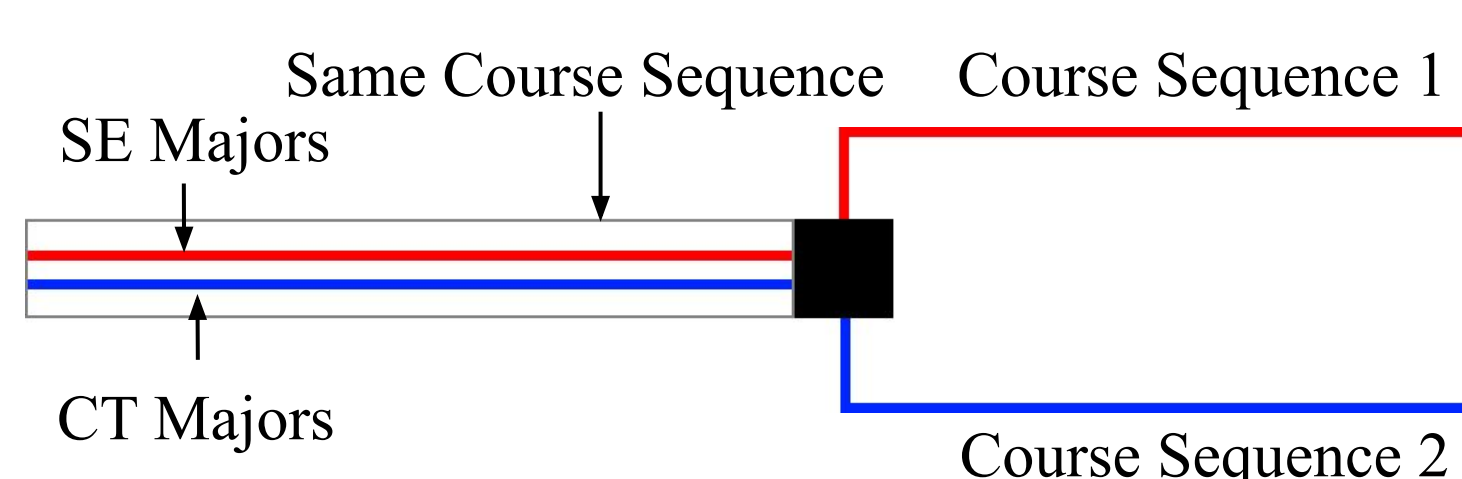


**Example:** What would have happened if a SE Major took sequence 2 instead of sequence 1?

## Experiment Design
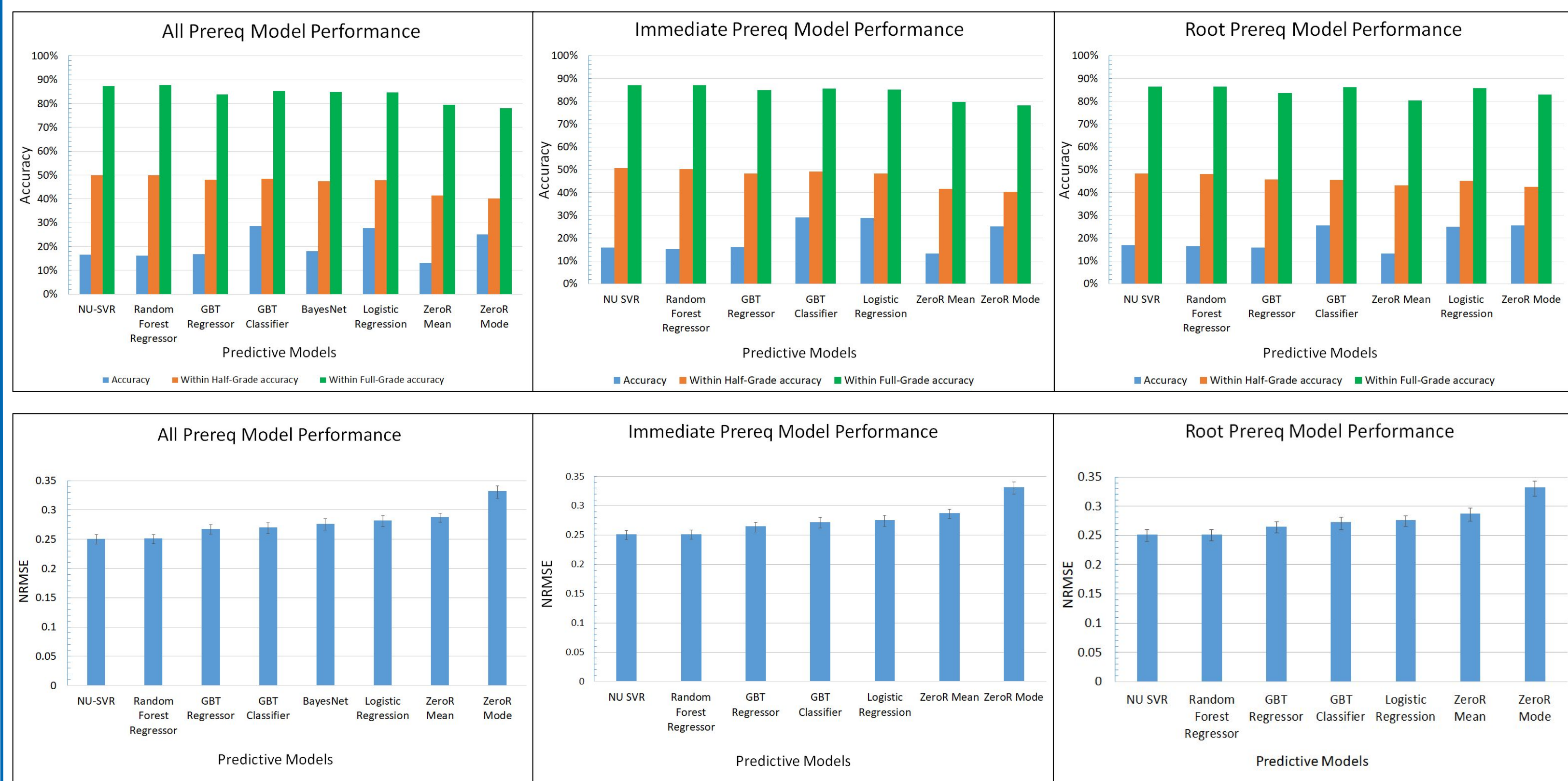


## Course Sequences



CIS & CT → CIS     SE → CT[30-59]
SE → CIS     SE
SE → CT[60-89]     CT

## Bayesian Network



Noisy Bayesian Network model structure with CS 2630 data

Prerequisite Node

| CS 2430 Grade | P |
|---|---|
| A | 0.18 |
| A- | 0.09 |
| B+ | 0.05 |
| B | 0.29 |
| ... | |
| F | 0.03 |

Auxiliary Node

| CS 2430 Grade | CS 2630 Grade | P |
|---|---|---|
| A | A | 0.58 |
| A | A- | 0.09 |
| A | B+ | 0.07 |
| A | B | 0.16 |
| ... | | |
| F | F | 0.56 |

Target Course Node

| CS 1130 Grade | CS 1430 Grade | CS 2430 Grade | CS 2630 Grade | P |
|---|---|---|---|---|
| A | A | A | A | 0.02 |
| A | A | A | A- | 0.15 |
| A | A | A | B+ | 0.26 |
| A | A | A | B | 0.23 |
| ... | | | | |
| F | F | F | F | 0.05 |
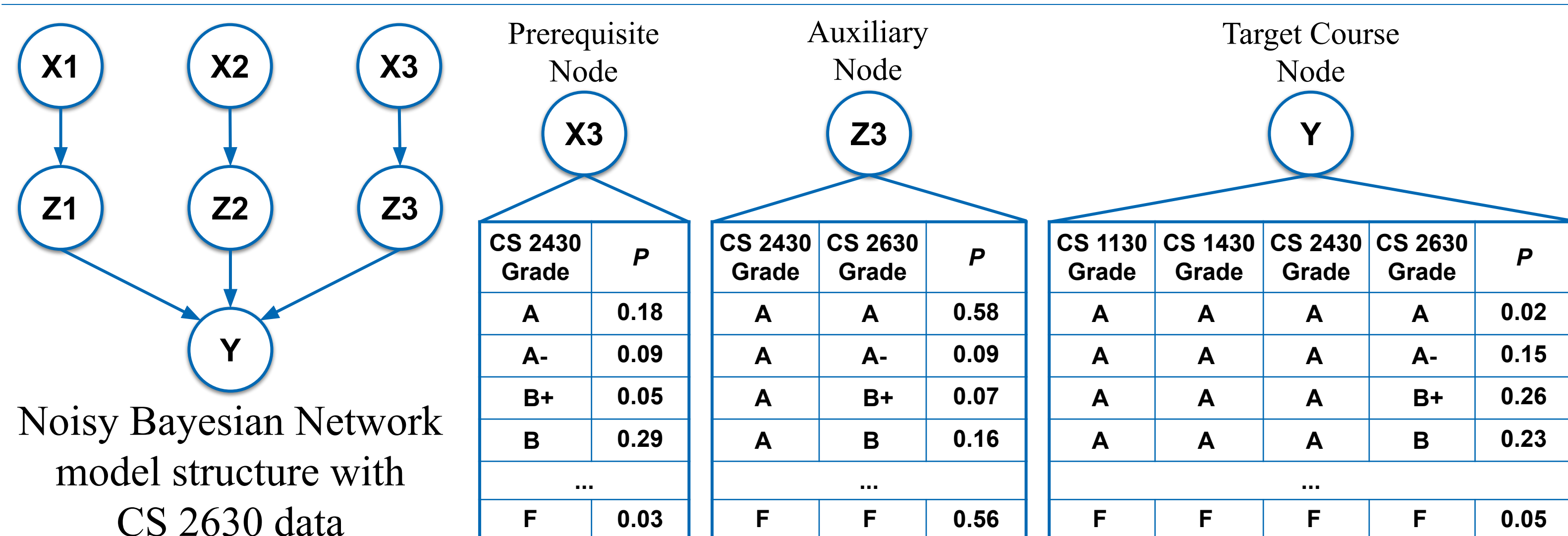
## Model Performances



## Counterfactual Inference Experiment Results



- We do not believe these results are accurate based on our expert knowledge and evidence of various forms of data shift
- **Covariate:** Shift in distributions of the independent variables
- **Prior Probability:** Shift in distributions of the target variable
- **Concept:** Shift in the relationships between the independent and target variables

*All dismissed students are also struggling*



| Major | AUC (Programming in C++) |
|---|---|
| CT → SE | 0.84 |
| CT → CIS | 0.86 |
| SE → CIS | 0.83 |

*A majority of struggling students weren't eligible to be used in our experiment, so they are not helpable*

*An example of test results for covariate shift with CS 1430 grades.  An AUC score ≥ 0.8 signifies that there is covariate shift*



*Example of concept shift in our dataset*        *Example of prior probability shift in our dataset*        *Example of covariate shift in our dataset*

## Future Work

- Retrain and rerun experiments accounting for all forms of data shift
- Extend the Bayesian Network model to a Dynamic Bayesian Network to account for the time differences between taking courses
- Test the external validity of the model with data from other programs and universities
- Create a recommender system that shows outcomes for possible course sequences to students/advisers.

## Conclusion

- Six common source sequences were identified in the CSSE curriculum
- Course grades can be predicted within a full letter grade with 85-90% accuracy
- Course grades can be predicted with root prerequisite course grades comparably well as immediate prerequisite course grades.
- Data shift was shown to occur and believed to have a large impact on the results. This may also be prevalent but not accounted for in existing course advising systems
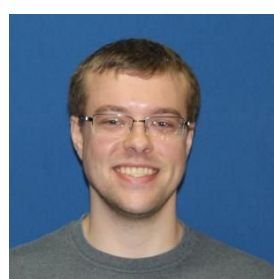
## Acknowledgements

## Authors

**Christopher Kott** (kottc@uwplatt.edu) is a senior at the University of Wisconsin - Platteville studying Software Engineering.

**Zachary Meyer** (meyerza@uwplatt.edu) is a senior at the University of Wisconsin - Platteville studying Computer Technology

**Lincoln Schroeder** (schroederli@uwplatt.edu) is a senior at the University of Wisconsin - Platteville studying Software Engineering

**Nicholas Tiede** (tieden@uwplatt.edu) is a junior at the University of Wisconsin - Platteville studying Software Engineering

**Dr. Douglas Selent** (selentd@uwplatt.edu) is an Assistant Professor in the Computer Science and Software Engineering department at the University of Wisconsin - Platteville. https://sites.google.com/site/dougselent/