Carnegie Mellon University
Master of Computational Data Science

Carnegie Mellon University
Language Technologies Institute

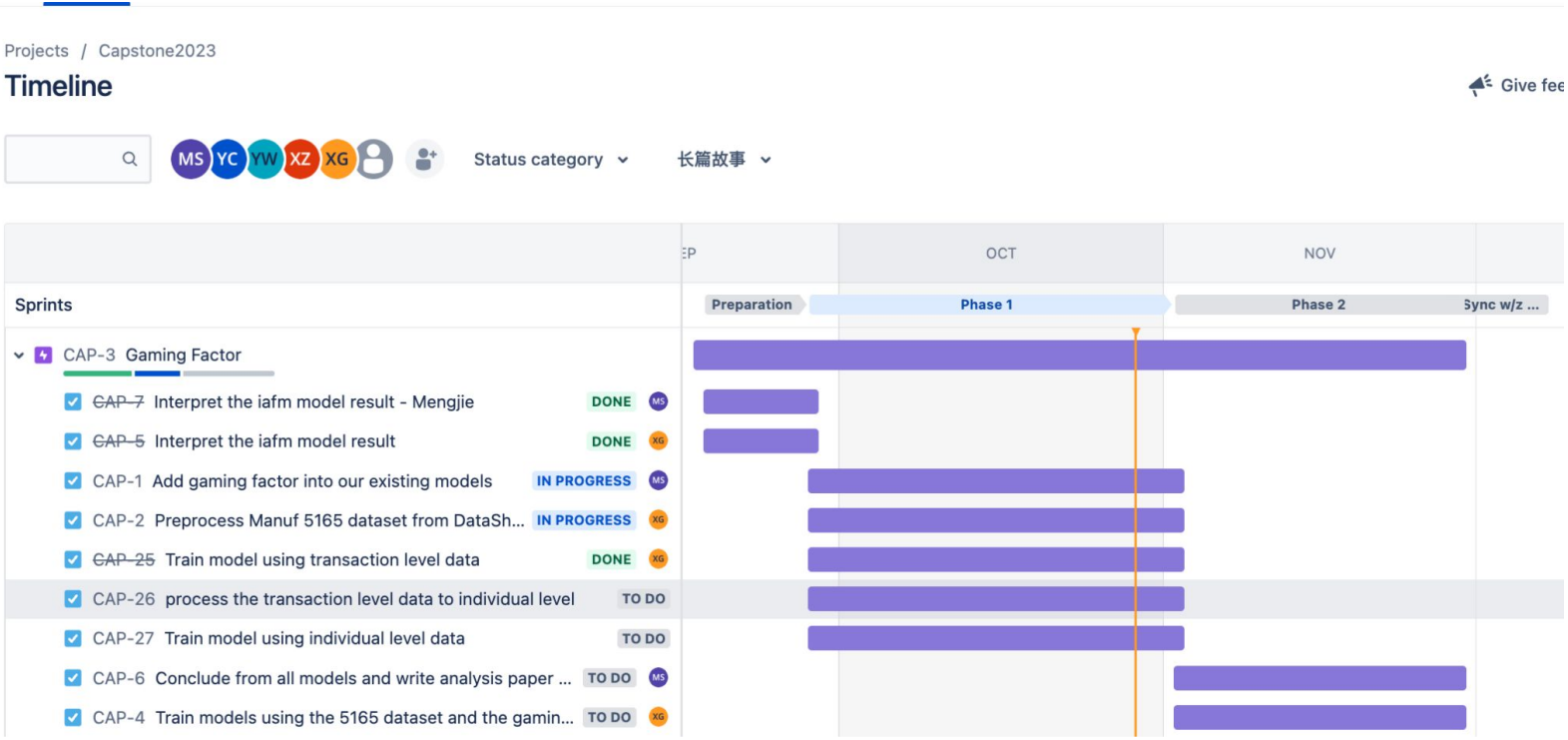
Predicting Learning Outcomes

Oct 30 Standup

Mengjie Shen, Xiaoyu Zhang, Xinyu Gu, Yizhou Chen, Yuchen Wang
Advisor: Ken Koedinger

11-632 (Fall 2023)
MCDS Capstone Course

Sub Group 1-Mengjie, Xinyu



Sub Group 1-Mengjie, Xinyu

Last week:

- Revisit dataset with mentors
  - joined two datasets to get the target features
  - merged the data from transaction-level to student step-level

This week:

- Define gaming and convert transaction-level gaming variables into individual-level gaming variables
- Train model using new datasets
- interpret the results & significance of features

Sub Group 2-Xiaoyu Zhang, Yizhou Chen, Yuchen Wang

Last week:

- Process the new dataset (ds5165) to align with the format of the Bernacki dataset
- Train the step-two linear regression model on the processed dataset to predict the student's post-test score with the individual's pre-test score and the iAFM model outcome averaged across all KCs

```
test_scores_with_prediction %>%
  lm(Posttest ~ PredAvgiAFM + Pretest, data = .) %>%
  summ()
```

```
MODEL INFO:
Observations: 129
Dependent Variable: Posttest
Type: OLS linear regression

MODEL FIT:
F(2,126) = 49.19, p = 0.00
R² = 0.44
Adj. R² = 0.43

Standard errors: OLS
------------------------------------------------
                    Est.   S.E.   t val.      p
----------------- ------ ------ -------- ------
(Intercept)         0.09   0.05     1.86   0.07
PredAvgiAFM         0.41   0.10     4.22   0.00
Pretest             0.46   0.09     5.34   0.00
------------------------------------------------
```

Sub Group 2-Xiaoyu Zhang, Yizhou Chen, Yuchen Wang

This week:

- Inspect model results with respect to the research question: using student parameter estimates derived from modeling the learning process data (ds5165), how well and how best can we predict student performance on data a post-test after the learning process and a pre-test before it.
- Try to incorporate individual student parameters (int_iAFM) into the step-two model on the new dataset
- Report and analyze the result regarding the research question

Sub Group 2-Xiaoyu Zhang, Yizhou Chen, Yuchen Wang