# EPARS: Early Prediction of At-risk Students with Online and Offline Learning Behaviors

Yu Yang[1*], Zhiyuan Wen[1], Jiannong Cao[1], Jiaxing Shen[1], Hongzhi Yin[2], and Xiaofang Zhou[2]

[1] Department of Computing, The Hong Kong Polytechnic University, HK SAR, China
[2] School of ITEE, The University of Queensland, Australia

* **Email:** csyyang@comp.polyu.edu.hk

# Outline

- Background & Motivations
- Problem Statement of At-risk Student Early Prediction
- Data Collection & Overview
- Challenges in At-risk Student Early Prediction
- Methodologies
  - Statistical analysis
  - Multi-scale Bag-of-Regularity
  - Social Homophily Embedding
- Experiments & Results
- Conclusion

# Background

- Students at risk (STAR) refer to students requiring temporary or ongoing intervention for achieving academic success[1].
  - Gradually fail to sustain their studies and then drop out
- Raising public concern of dropout, depression, suicide etc.
- Diverse factors cause students being at-risk.
  - Poor academic performance
  - Family problems
  - Financial stress
  - Social barriers
  - …

[1] Richardson, V.: At-risk student intervention implementation guide. The Education and Economic Development Coordinating Council At-Risk Student Committee p. 18 (2005)

# Motivations

- Early prediction of STAR offer the opportunity to timely intervene.
- University usually identifies STAR by their academic performance.
  - Too late for interventions.
- Existing works predict STAR from either online or offline learning behaviors.
  - Hardly capture the whole learning processes in a comprehensive manner.
  - Unsatisfactory accuracy in STAR early prediction.

# Problem Statement

- STAR are students whose average Grade Point Average (GPA) is below 2.0 in a semester.
  - When a student has a GPA below 2.0, he/she will be put on academic probation in the following semester.
  - If a student cannot pull his/her GPA up to 2.0 or above in the semester, he/she will be dropped out.
- Problem formulation of STAR early prediction

**Given:**
- Students' click operations in the Blackboard (online learning traces)
- Students' library check-in records (offline learning traces)

**Objective:** Identify STAR as accurate and early as possible in a semester

# Data Collection & Overview

- Data Collection
  - Click-stream data with timestamps in the Blackboard
  - Library check-in logs
  - GPA
- Data Scope
  - All 15,503 undergraduate students in the whole university
  - 2016 to 2017 academic year

Table 1. Data Overview.

| | Semester 1 | | Semester 2 | |
|---|---|---|---|---|
| | STAR | Other Std | STAR | Other Std |
| Population | 391 | 15,112 | 225 | 15,278 |
| # click-stream logs in LMS | 2,225,605 | 95,949,014 | 1,019,134 | 70,874,428 |
| Avg. # click-stream logs | 5,692.0844 | 6,349.1936 | 4,529.4844 | 4,638.9860 |
| Avg. # click-stream logs in first 2 weeks | 301.4041 | 399.9502 | 243.0400 | 284.4368 |
| Avg. # click-stream logs in last 2 weeks | 526.6522 | 545.4346 | 336.9133 | 304.7331 |
| # library check-in | 14,045 | 636,353 | 6,245 | 517,557 |
| Avg. # library check-in | 35.9207 | 42.1091 | 27.7556 | 33.8760 |
| Avg. # library check-in in first 2 weeks | 1.7877 | 2.3303 | 1.3889 | 1.8424 |
| Avg. # library check-in in last 2 weeks | 2.9834 | 3.3760 | 2.3444 | 2.4547 |

# Challenges

- Data density imbalance
  - Offline learning records (library check-in) are much sparser than online learning traces (click-stream traces in the Blackboard).
  - The overall behavior representation will be easily dominated by the online learning behavior in fusion.
- Data insufficiency
  - Students, especially STAR, are usually inactive at the beginning of a semester.
  - The behavior traces are far from enough for accurate early prediction of STAR.
- Label imbalance
  - The number of STAR is far less than that of normal students.
  - STAR prediction is an extreme label-imbalance classification problem.

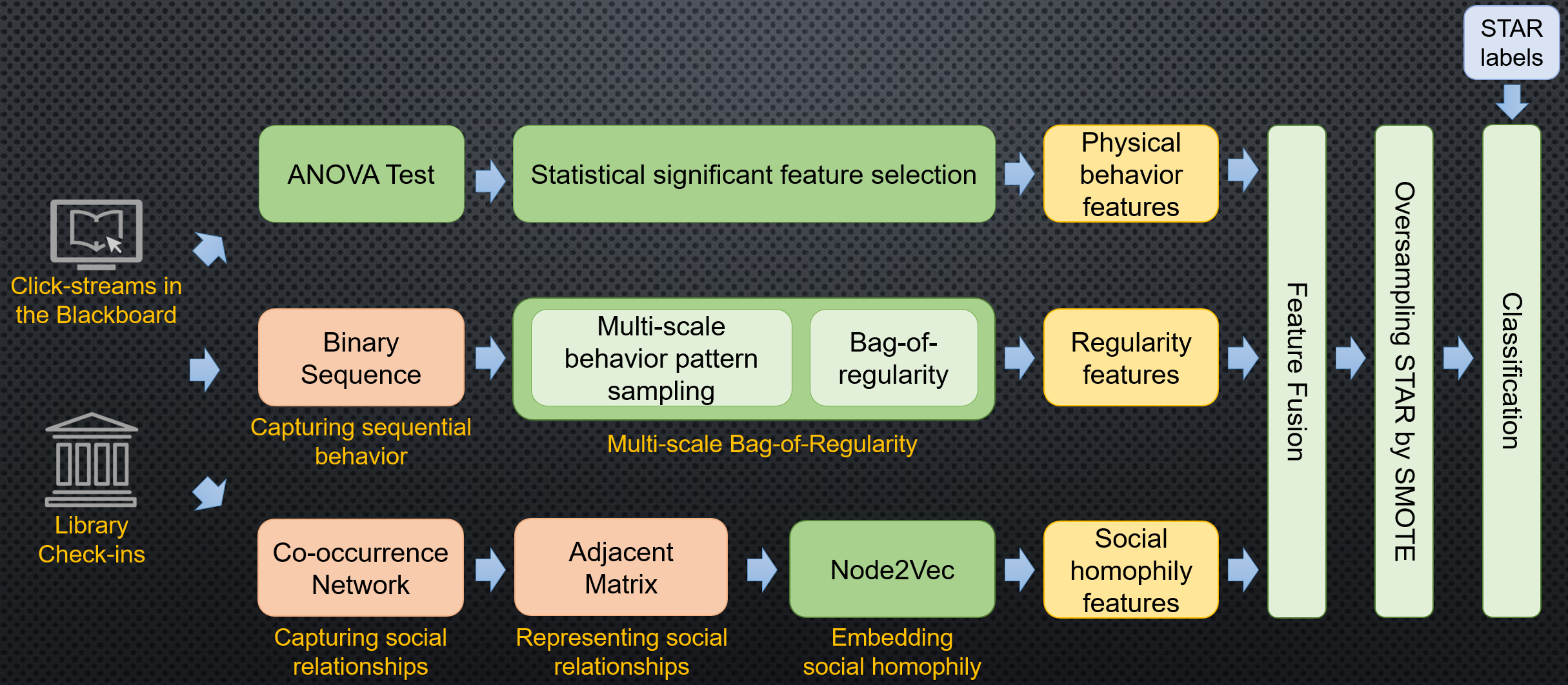
# Observations

- Study routines
  - Good students usually follow their study routines periodically and show clear regularities of learning patterns.
  - Study routines of STAR are disorganized leading to irregular learning patterns.

- Social homophily
  - Students tend to have social tie with others who are similar to them.
  - At-risk students had more dropout friends[2].

[2] Ellenbogen, S., Chamberland, C.: The peer relations of dropouts: a comparative study of at-risk and not at-risk youths. Journal of adolescence 20(4), pp. 355-367 (1997)

# Framework of EPARS



# Statistical Analysis by ANOVA

- Findings from the ANOVA test
  - STAR use the Blackboard less than the normal students
  - STAR check the announcement and lectures' information more than normal students
  - STAR go to the library less than the normal students at the beginning of a semester
  - STAR prefer more to go to the library after business hours

**Table 2.** Results of the ANOVA test.

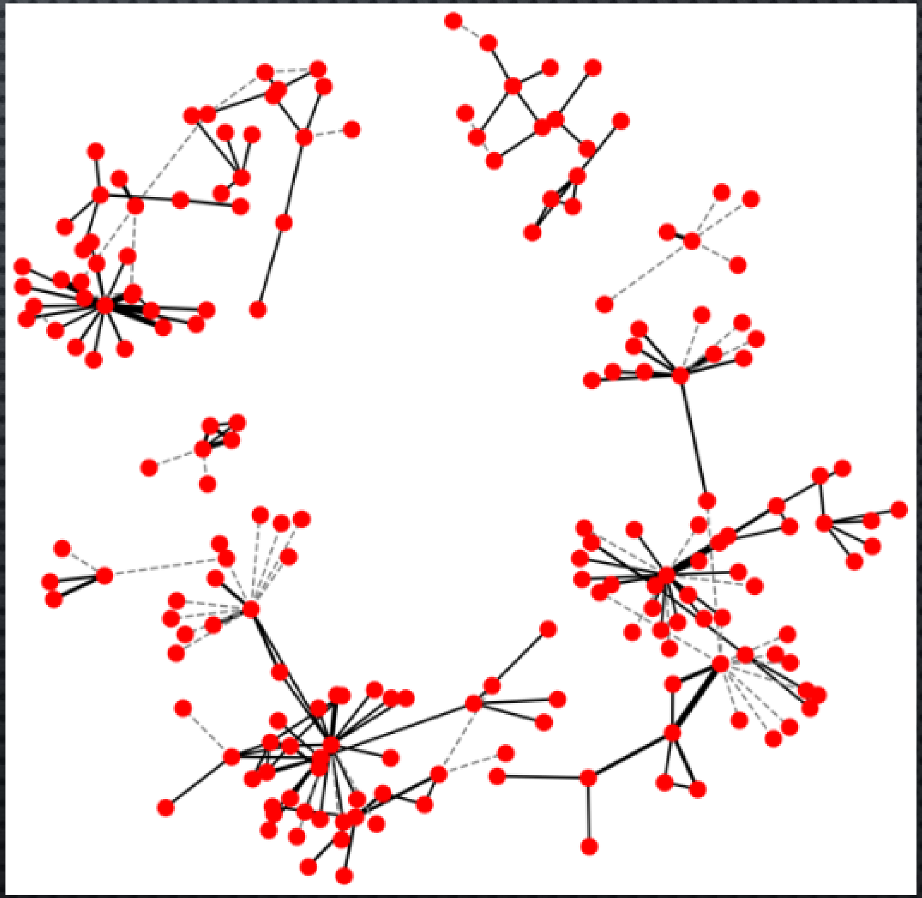
| Features | P-value | F-value | Mean STAR | Mean Others |
|---|---|---|---|---|
| # LMS Login | 0.0020 | 9.5112 | 127.4987 | 144.8043 |
| # LMS Logout | 0.0000 | 34.5301 | 8.9318 | 20.1348 |
| # Check announcement | 0.0158 | 5.8311 | 41.4436 | 36.8361 |
| # Course access | 0.7328 | 0.1165 | 4.2677 | 4.5667 |
| # Grade center access | 0.7694 | 0.0859 | 10.5486 | 10.2108 |
| # Discussion board access | 0.0020 | 9.5951 | 11.7979 | 19.2444 |
| # Group access | 0.0209 | 5.3385 | 13.2782 | 20.1268 |
| # Check personal info | 0.0000 | 16.7953 | 0.2283 | 1.6585 |
| # Check lecturer info | 0.0000 | 106.1638 | 9.7297 | 5.5440 |
| # Journal page access | 0.0199 | 5.4191 | 0.2283 | 1.6585 |
| # Lib check-in | 0.0700 | 3.2829 | 42.8163 | 47.3589 |
| # Lib check-in in the morning | 0.0001 | 14.7133 | 7.0367 | 9.4206 |
| # Lib check-in in the afternoon | 0.0023 | 9.3196 | 27.0604 | 31.9419 |
| # Lib check-in after midnight | 0.0000 | 43.9327 | 4.0105 | 1.6927 |
| # Lib check-in before exam months | 0.0123 | 6.2740 | 33.9265 | 39.0143 |
| # Lib check-in at the first month | 0.0004 | 12.5447 | 8.4724 | 10.6052 |

# Multi-scale Bag-of-Regularity

- Construct a binary sequence from students' sequential behavior traces
  - Mark as 1 if the learning behavior appears, i.e. go to the library, log-in the LMS
- Multi-scale behavior pattern sampling
  - Subsequences of length $\ell = 2 + (s - 1) \times z$ centered on nonzero elements
    - $s \in \{1,2, \ldots, S\}$ is the scale.
    - $z$ is the step-size between scales.
  - All zero subsequences are excluded for overcoming the sparsity problems
- Bag-of-Regularity
  - Treat all possible behavior patterns excluding all-zeros one as a bag.
  - Count the number of occurrences of every sampled behavior pattern.

# Embedding Social Homophily

- Modeling social relationship by constructing a co-occurrence network from the library check-ins
  - Intuitions: If students are friends, they are more likely to learn together.
  - Co-occurrence: The time difference of the library check-in between two students is less than a threshold $\delta$.
  - Distinguish familiar strangers: # co-occurrence in the library is more than a threshold $\sigma$.
- Embedding social homophily by Node2Vec
  - Constrains: The features of students who have similar social connections should be close

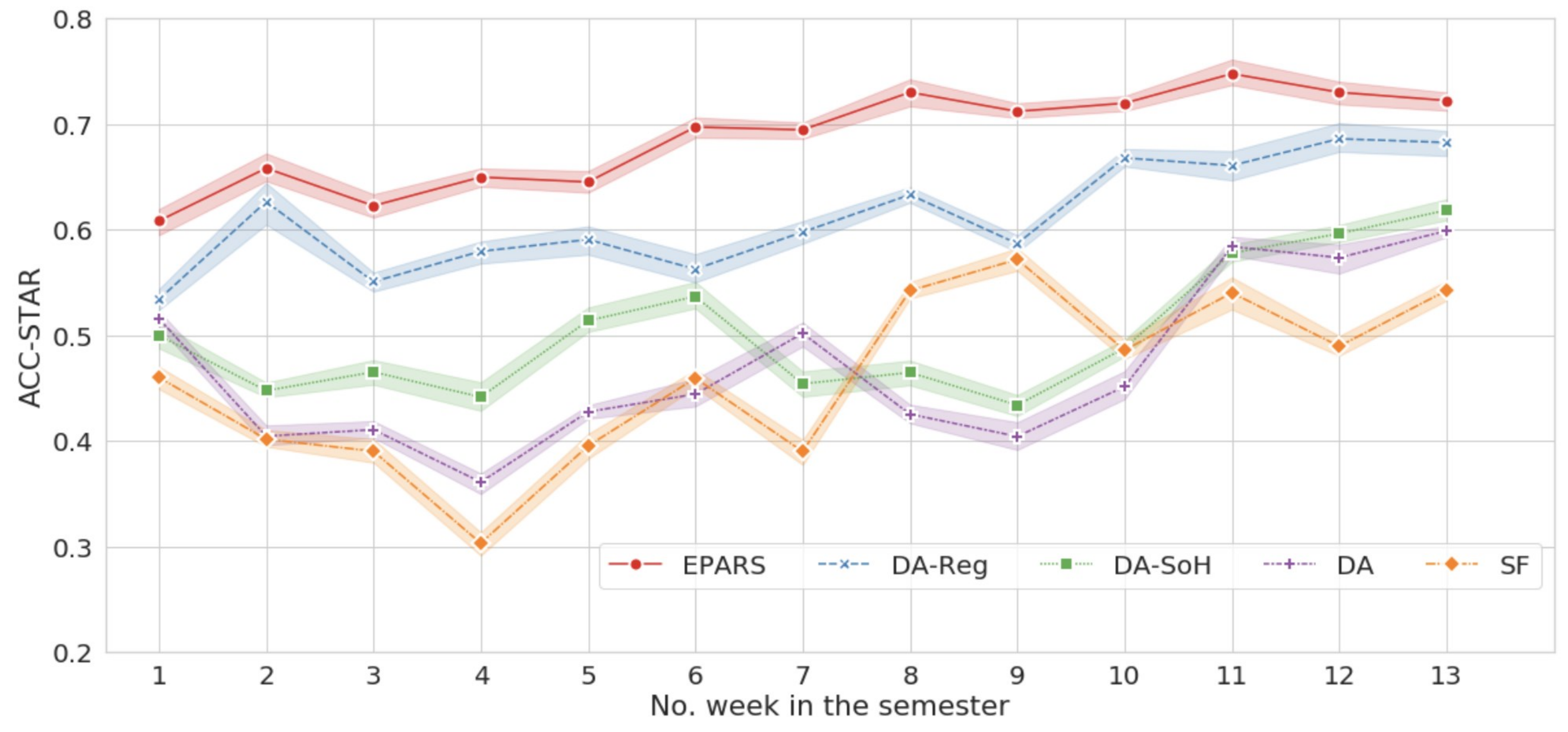


A part of the constructed co-occurrence network with $\sigma = 5$.

# Experiment Protocol

- Experiment Setting
  - Predicting STAR at the end of every week in the semester using the data collected from the beginning of the semester to time making prediction.
  - Under 5-fold cross-validation setting and repeat 10 times.
  - Report the average results obtained by the Gradient Boosting Decision Tree.
- Evaluation metrics
  - AUC: Areas under the ROC curve
  - ACC-STAR: The amount of true positive predictions divided by the total number of STAR
- Baselines
  - SF: statistical significant features by ANOVA testing
  - DA: SF + data augmentation
  - DA-SoH: SF + data augmentation + social homophily embeddings
  - DA-Reg: SF + data augmentation + regularity features

# Results of STAR Early Prediction




| ACC-STAR | SF | DA | DA-SoH | DA-Reg | EPARS |
|---|---|---|---|---|---|
| First week | 0.4474 | 0.5263 | 0.5053 | 0.5395 | **0.6184** |
| Last week | 0.5395 | 0.6079 | 0.6184 | 0.6842 | **0.7237** |

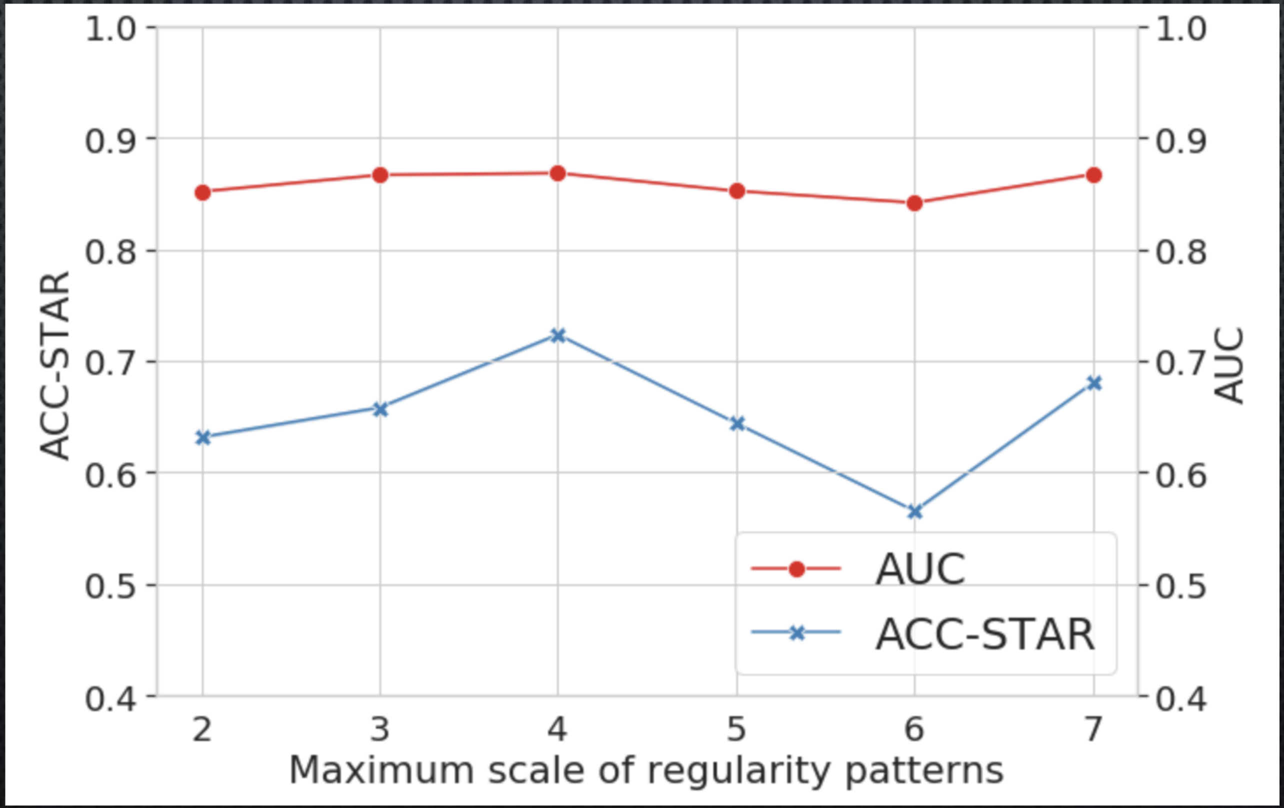
# Evaluation of Data Augmentation

- SMOTE achieves the best STAR prediction accuracy
  - Increases the number of minority samples
  - Enriches the diversity of the training set

| | # STAR after DA per fold | # Normal Std. after DA per fold | AUC | ACC-STAR |
|---|---|---|---|---|
| No DA | 305 | 11295 | 0.8342 | 0.5526 |
| Random Under-sampling | 305 | 305 | 0.8211 | 0.5316 |
| Random Over-sampling | 11295 | 11295 | 0.8458 | 0.5645 |
| SMOTE | 11295 | 11295 | **0.8684** | **0.7237** |

# Sensitivity of Maximum Scale

- EPARS achieves the best performance when maximum scale $S = 4$.
  - Regularity patterns of the scale 5 to 7 can be synthesized by the scale of 2 to 4.
  - Regularity features will dramatically become sparse when $S > 4$.



# Co-occurrence Parameters Sensitivity

- Testing time difference threshold $\delta$ for determining co-occurrence

Results of Testing Time Difference Threshold

| $\delta$ | Ave # edges per week | AUC | ACC-STAR |
|---|---|---|---|
| 10 seconds | 14263 | **0.8699** | 0.5921 |
| 30 seconds | 39386 | 0.8684 | **0.7273** |
| 60 seconds | 77318 | 0.8576 | 0.6316 |

- Testing linking threshold $\sigma$ for filtering familiar strangers

Results of Testing Linking Threshold

| $\sigma$ | AUC | ACC-STAR |
|---|---|---|
| 2 times | **0.8684** | **0.7237** |
| 3 times | 0.8615 | 0.6184 |
| 4 times | 0.8554 | 0.5658 |
| 5 times | 0.8122 | 0.5395 |

# Conclusion

- A novel algorithm EPARS for early predicting STAR.
  - Extract students' learning regularity patterns and social homophily from online and offline learning behaviors.
- A multi-scale bag-of-regularity method to extract regularity features from sequential learning behaviors.
  - Robust for sparse data
- Embedding social homophily from a co-occurrence network constructed from library check-ins.
  - Supplement the lack of behavior traces for STAR
- EPARS is accurate in STAR early prediction
  - 14.62% ~ 38.22% accuracy improvement to the baselines even in the first week of a semester

thank you!