EDUCATIONAL DATA MINING

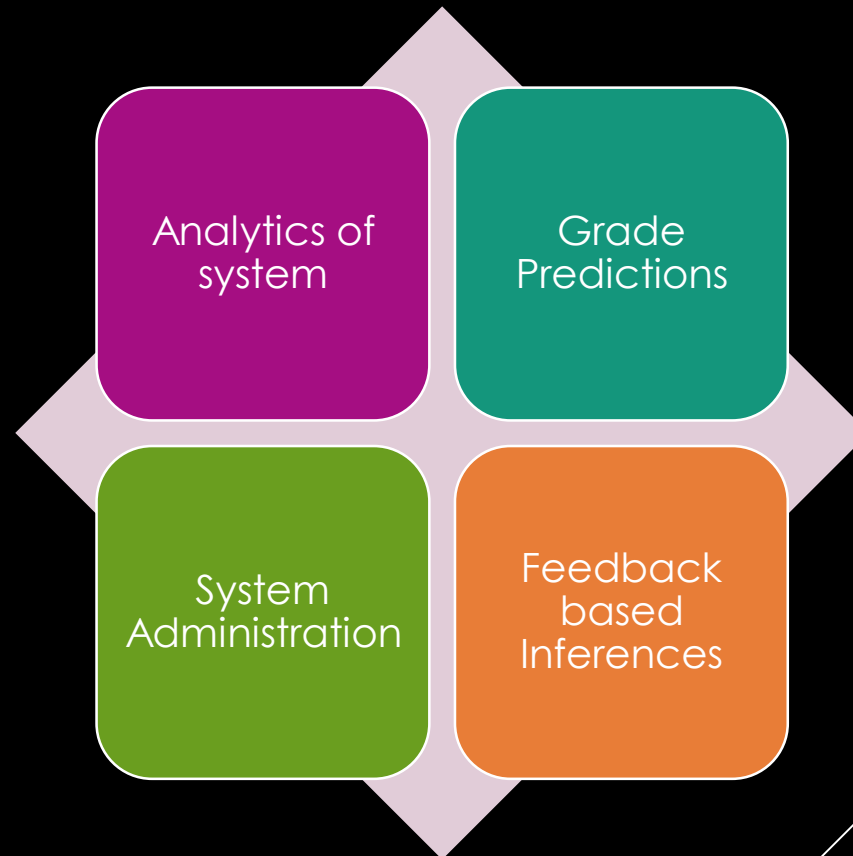# IBM DATA SCIENCE CAPSTONE PROJECT

* EDM (Educational Data Mining) means mining insights from Educational Data.

LITERATURE REVIEW

- Recent advancements in the domain

LITERATURE REVIEW

- Educational Scenario around the world

United States	India
5,300 Colleges and Universities	60,000 Colleges and Universities
Uniformity in Educational System and Programs	Non-Uniform System and Programs
Lesser Financial Dependencies	More Financial Dependencies
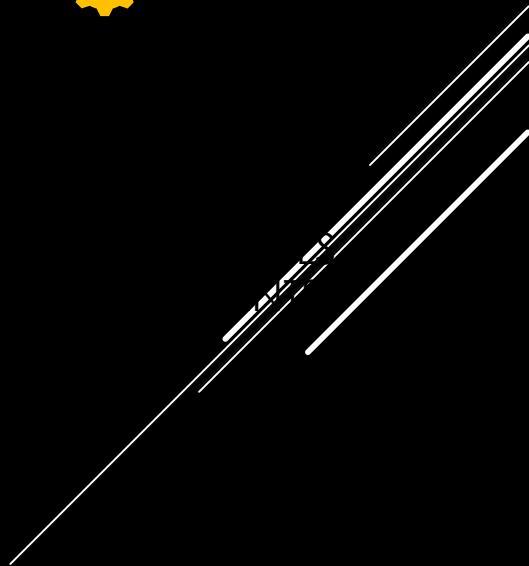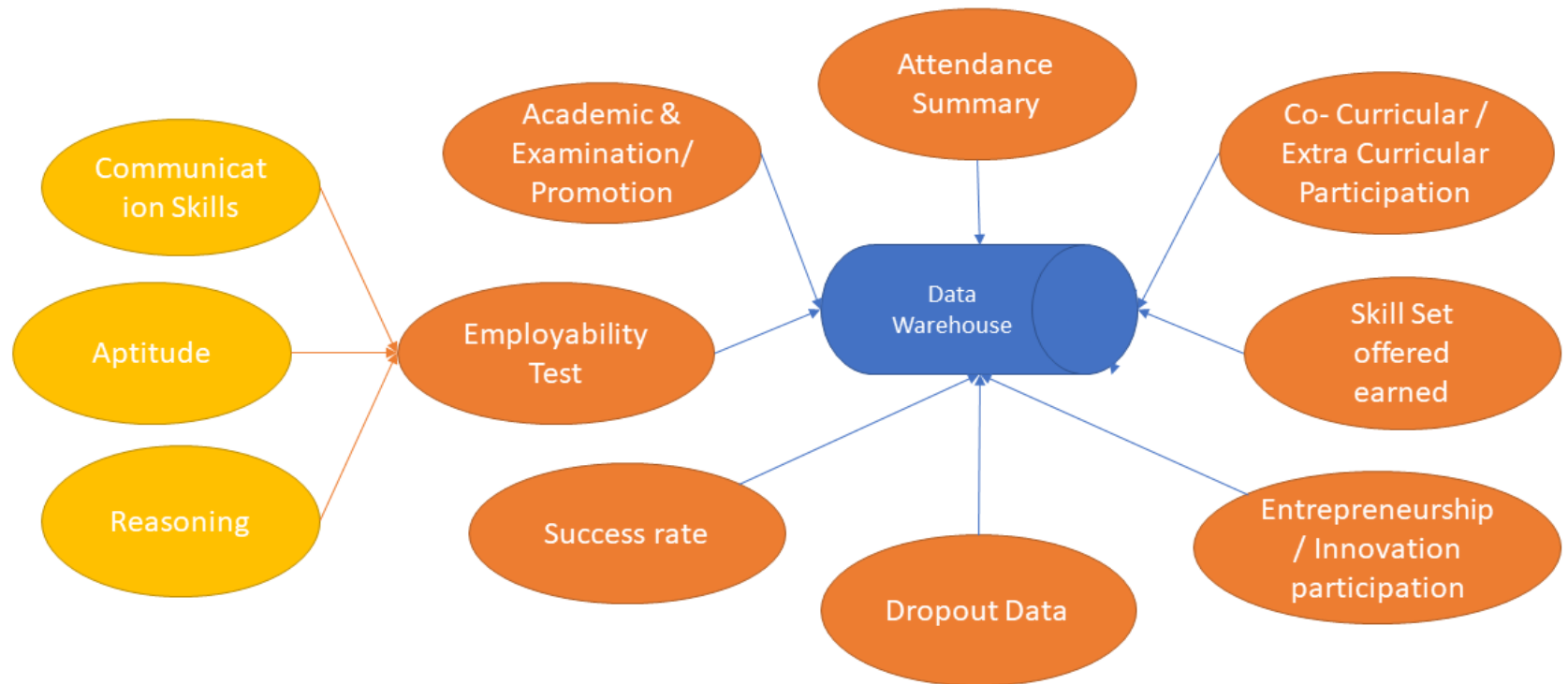
Early drop out prediction

Future Performance Prediction

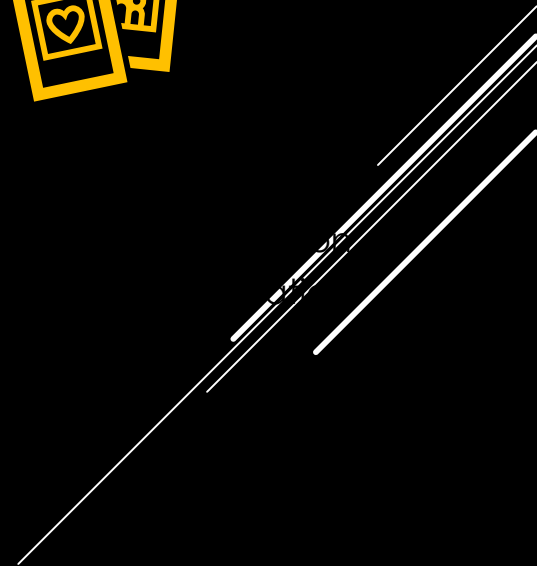Early measures towards Performance Improvement based on prediction
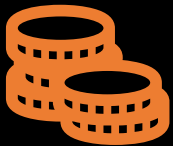
Domain Modelling

Student Feedback based Inferences & Changes in Policies

Graduate Admission Prediction & Course Suggestion Module

- Assists Institutes in potential intake of candidates.

Learning habits of student generally remains same

EARLY MEASURES TOWARDS PERFORMANCE IMPROVEMENT BASED ON PREDICTION

- Extracting skill set of candidates from Data Warehouse

Classification into extra-curricular or co-curricular domain	Aim to suggest cluster to obtain higher productivity

STUDENT FEEDBACK BASED INFERENCES & CHANGES IN POLICIES

Study collective class performance

Teachers assessment for management

Change in policies

GRADUATE ADMISSION PREDICTION & COURSE SUGGESTION MODULE

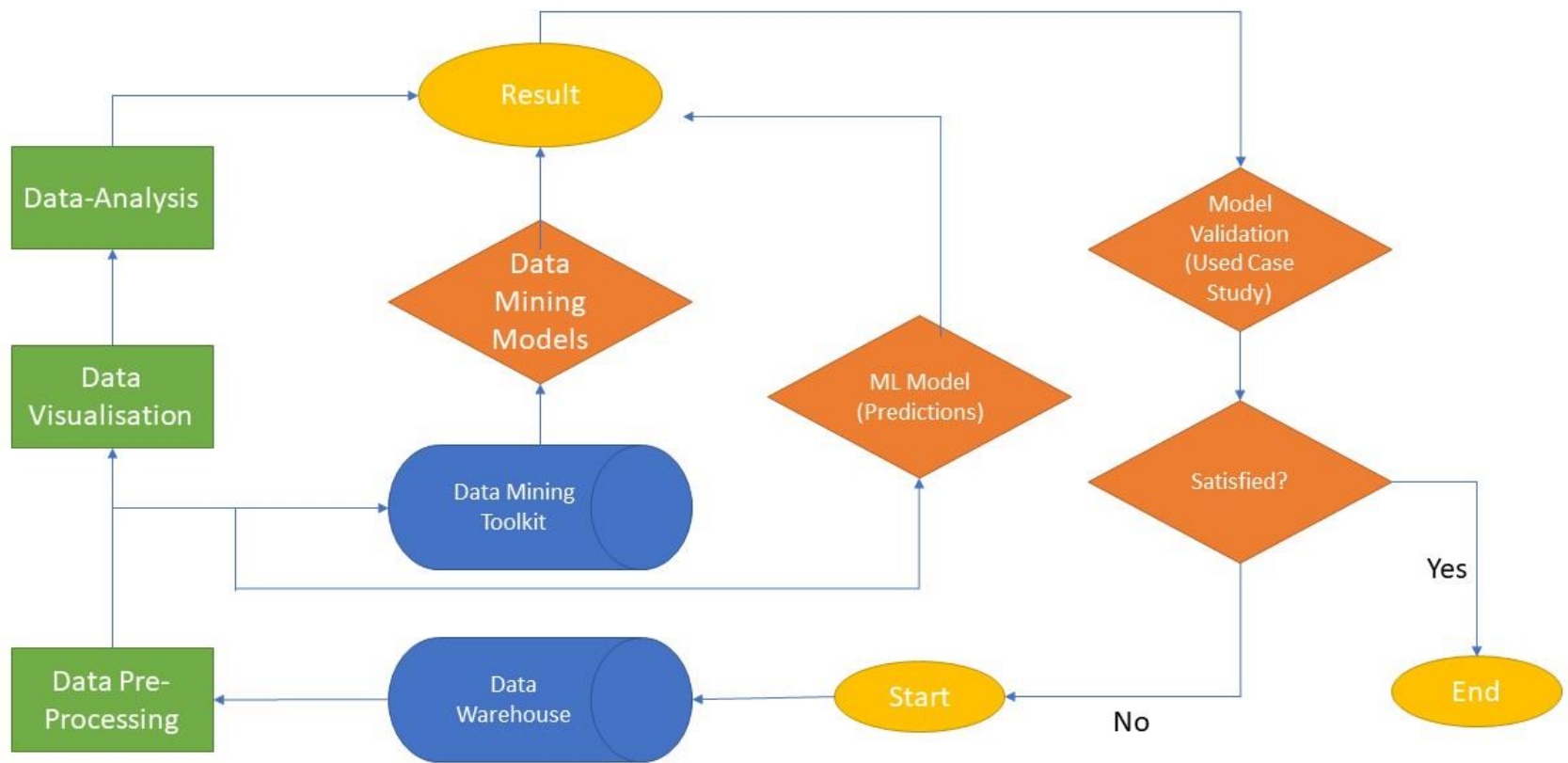# GENERAL METHODOLOGY

CHALLENGES

Diversity in Educational System

Diversity of students

CONCLUSION

New Sub-domains Identified under EDM

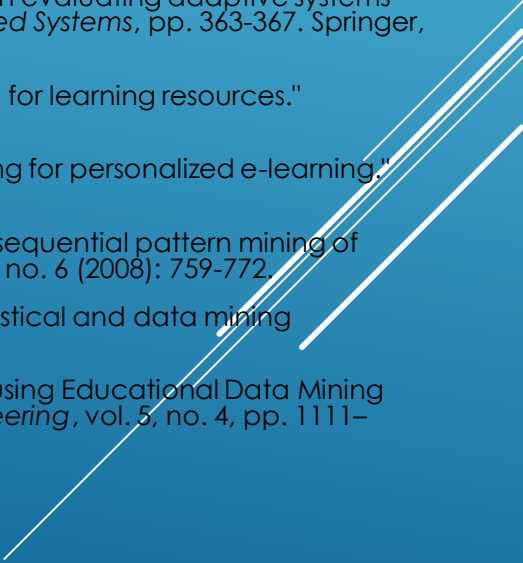Different Educational Scenario likely to affect outcomes.

Diversity in population & educational system likely to affect.

Socio-economic status has higher correlations with results.

REFERENCES

1. Ha, Sung Ho, Sung Min Bae, and Sang Chan Park. "Web mining for distance education." In *Proceedings of the 2000 IEEE International Conference on the Management of Innovation and Technology. ICMIT 2000.'Management in the 21st Century'(Cat. No. 00EX457)*, vol. 2, pp. 715-719. IEEE, 2000

2. Sheard, Judy, Jason Ceddia, John Hurst, and Juhani Tuovinen. "Inferring student learning behaviour from website interactions: A usage analysis." *Education and Information Technologies* 8, no. 3 (2003): 245-266.

3. Ritter, S., Harris, T., Nixon, T., Dickison, D., Murray, R., Towle, B. (2009). Reducing the knowledge tracing space. In International Conference on Educational Data Mining, Cordoba, Spain, 151-160.

4. Shen, Ruimin, Fan Yang, and Peng Han. "Data analysis center based on e-learning platform." In *The Internet Challenge: Technology and Applications*, pp. 19-28. Springer, Dordrecht, 2002.

5. Tane, Julien, Christoph Schmitz, and Gerd Stumme. "Semantic resource management for the web: an e-learning application." In *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters*, pp. 1-10. ACM, 2004.

6. Arruabarrena, Rosa, Tomás A. Pérez, Javier López-Cuadrado, J. Gutiérrez, and J. A. Vadillo. "On evaluating adaptive systems for education." In *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, pp. 363-367. Springer, Berlin, Heidelberg, 2002.

7. Cummins, Dean, Kalina Yacef, and Irena Koprinska. "A sequence based recommender system for learning resources." *Australian Journal of Intelligent Information Processing Systems* 9, no. 2 (2006): 49-57.

8. Köck, Mirjam, and Alexandros Paramythis. "Activity sequence modelling and dynamic clustering for personalized e-learning." *User Modeling and User-Adapted Interaction* 21, no. 1-2 (2011): 51-97.

9. Perera, Dilhan, Judy Kay, Irena Koprinska, Kalina Yacef, and Osmar R. Zaïane. "Clustering and sequential pattern mining of online collaborative learning data." *IEEE Transactions on Knowledge and Data Engineering* 21, no. 6 (2008): 759-772.

10. Ramesh, V. A. M. A. N. A. N., P. Parkavi, and K. Ramar. "Predicting student performance: a statistical and data mining approach." *International journal of computer applications* 63, no. 8 (2013): 35-39.

11. S. Gole, P. Bajaj, and A. Thomas, "Prediction of Dropout Students from Engineering Education using Educational Data Mining (EDM)," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 5, no. 4, pp. 1111–1118.

12. HRD to increase nearly 25 pc seats in varsities to implement 10 pc quota for poor in gen category, Economic Times, 16 January 2019.

13. Brusilovsky, P., & Peylo, C. (2003). Adaptive and intelligent web-based educational systems. International Journal of Artificial Intelligence in Education, 13, 156–169.

14. Zaiane, Osmar. "Web usage mining for a better web-based learning environment." (2001).

15. Sundar, P. P. "A comparative study for predicting students academic performance using Bayesian network classifiers." *IOSR Journal of Engineering (IOSRJEN) e-ISSN* (2013): 2250-3021.

16. Agarwal, Pawan. "Higher education in India: Growth, concerns and change agenda." *Higher Education Quarterly* 61, no. 2 (2007): 197-207.

17. Gupta, Deepti, and Navneet Gupta. "Higher education in India: structure, statistics and challenges." *Journal of education and Practice* 3, no. 2 (2012).

18. Blom, Andreas, and Hiroshi Saeki. *Employability and skill set of newly graduated engineers in India*. The World Bank, 2011.

19. Saa, Amjad Abu. "Educational data mining & students' performance prediction." *International Journal of Advanced Computer Science and Applications* 7, no. 5 (2016): 212-220.

20. Pandey, U.K. and Pal, S., 2011. Data Mining: A prediction of performer or underperformer using classification. (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (2), 2011, 686- 690.

21. Hu, Ya-Han, Chia-Lun Lo, and Sheng-Pao Shih. "Developing early warning systems to predict students' online learning performance." *Computers in Human Behavior* 36 (2014): 469-478.

22. Heppen, Jessica B., and Susan Bowles Therriault. "Developing Early Warning Systems to Identify Potential High School Dropouts. Issue Brief." *National High School Center* (2008).