Sequence based course recommender for personalised curriculum planning

Chris Wong cwon7609@uni.sydney.edu.au
School of Information Technology

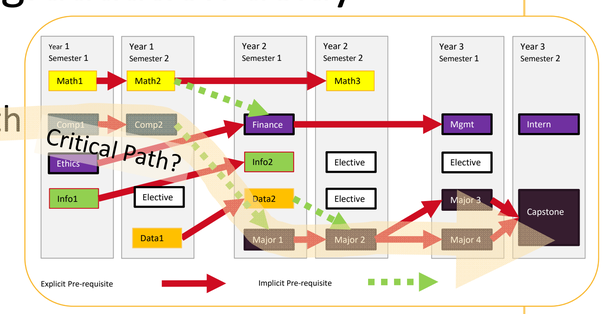
Motivation

Help students navigate their degrees

- Influence student success via course advice
- Minimise degree completion time
- Maximise results - Grade Point Average (GPA)
- This study focuses on institutional (structural/operational) factors rather than personal background factors
  ….e.g. preparedness, medical, employment, social/friendship, financial…

Critical paths could result in graduation delay

- Not just poor choices
  …Even with pre-defined pathways
- Failure or deferment on critical path could lead to cascading delays
  … but not all critical paths are explicitly known…



Other issues affecting choice:

- Structural: Sequencing, concurrency (*compulsory, pre-requisites, …*)
- Operational: constraints (*course availability, scheduling, class quotas …*)
- Individual: context (*intended career, jobs, interests , accreditation …*)

➔ increase complexity of decision making

Advisers are key sources of pathways

- Sources are academic advisors, peers, online, friends, social media
- Using their Implicit/tacit knowledge?
- Not scalable, inconsistent, outdated or inaccurate advice.

Related Work

Existing systems

- Course catalogues, handbooks/calendars, study plan templates, enrolment systems, integrated with student management systems

Machine learning for educational data science

- Educational data mining:
  … used for analysis and prediction [3,6](grades, failure risk)
- Educational recommenders[8] becoming common:
  … for career choice/degree/major, learning objects[1] using collaborative filtering, content based approaches or association mining[2].

Deep Learning (Recurrent Neural Networks)

- RNN now being used for sequence prediction[4], learning analytics[9], student performance prediction[5] & contextual recommendations[7]

Data

10 years of student records

- 2.1 million transcript results[11]
- 30 degrees, 14 majors, 400+ courses, 72000 graduation records
- ➔ use subset for experiment – Engineering and IT records

Preparation

- Data wrangling, cleansing (*anomalies, missing data, …)*
- Normalisation (*due to part-time, absences, double degrees, transfers, credit transfer …*)
- k-anonymised, de-identified
- Human ethics approval[12]!

Attributes to be used:

- Degree/major, course, year/semester, grades/marks, classification information e.g. gender, status
- Courses *may* be annotated by topics & learning objectives

Methodology

Exploratory Data Analysis

- Use Machine Learning – clustering (k-means) , predictive modelling (regression, decision tree, neutral network, random forest) → detect patterns & make observations
- Analyse by cohorts e.g. engineering majors, IT

Develop Baseline Recommender

- Results are used for implementing classical recommender → becomes “baseline” for evaluation

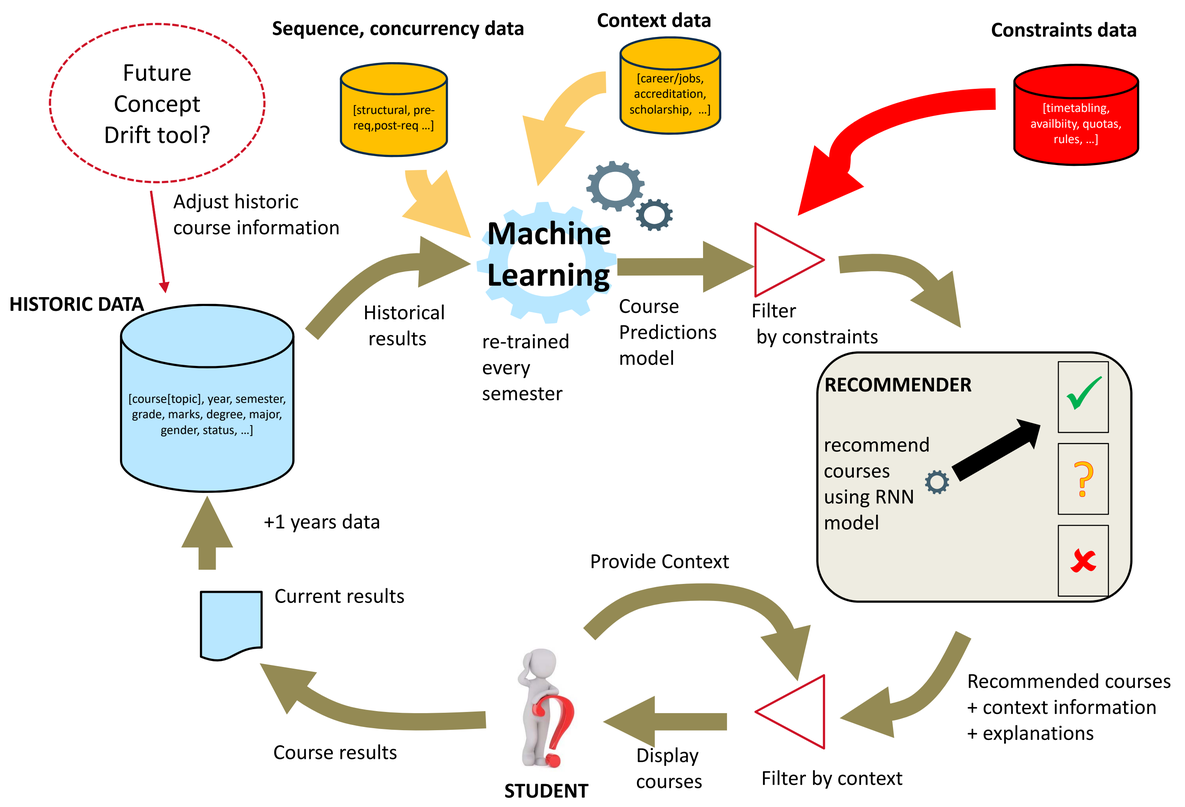
Prototype & test Initial deep learning recommender

- Train & tune Recurrent Neural Network & compare to baseline
  (objective function: hybrid of duration and GPA, add sequence , context and constraint filters)
- Evaluate – using split (90:10) and 10-fold cross validation
- Prototype user interface options (including HCI, formatting, explainability, warnings)

Experiment & evaluation

(see below..)

Proposed Solution



Experiment

Pilot prototype on focus group

- Select focus group users – pre& post pilot qualitative/quantitative surveys, workshop on prototype. Ensure appropriate mix of 1st, 2nd, 3rd & final year students and faculty/admin users.
- These students will use prototype to choose courses.
- Compare baseline (i.e. simulation of other peer students)
- Adjust database for concept drift as courses may have changed.
- Attempt multi-semester → retrain deep learning model based on student results.

Issues

Ethical

- Monitor focus group users – do no harm!

User interface

- How to design and present recommendations?
- How to provide explanations of recommendations?

Concept Drift[10] automation

- Needs further research – e.g. topic mining course descriptions/outlines, develop heuristics, adjust historic data?

References: **1**. Burke, R., Abdollahpouri, H.: *Educational Recommendation with Multiple Stakeholders*. In: 2016 IEEE/WIC/ACM International Conference on Web Intelligence Workshops (WIW). pp. 62–63 IEEE (2016) **2**. Bydžovská, H.: *Course Enrollment Recommender System*. Proc. 9th Int. Conf. Educ. Data Min. 1, 312–317 (2016). **3**. Bydžovská, H., Popelinsky, L.: *Predicting student performance in higher education*. Proc. - Int. Work. Database Expert Syst. Appl. DEXA. 141–145 (2013). **4**. Devooght, R., Bersini, H.: *Long and Short-Term Recommendations with Recurrent Neural Networks*. In: Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization - UMAP ’17. pp. 13–21 (2017). **5**. Okubo, F. et al.: *A neural network approach for students’ performance prediction*. In: Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK ’17. pp. 598–599 ACM Press, New York, New York, USA (2017). **6**. Polyzou, A., Karypis, G.: *Grade prediction with course and student specific models*. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 9651, 3, 89–101 (2016). **7**. Smirnova, E., Vasile, F.: *Contextual Sequence Modeling for Recommendation with Recurrent Neural Networks*. In: 2nd workshop on Deep Learning for Recommender Systems (DLRS 2017). pp. 2–9 ACM Press, New York, New York, USA (2017). **8**. Sweeney, M. et al.: *Next-Term Student Performance Prediction: A Recommender Systems Approach*. In: International conference on Educational Data Mining. (2016). **9**. Tang, S. et al.: *Deep Neural Networks and How They Apply to Sequential Education Data*. In: Proceedings of the Third (2016) ACM Conference on Learning @ Scale - L@S ’16. pp. 321–324 ACM Press, New York, New York, USA (2016). **10**. Žliobaitė, I. et al.: *An Overview of Concept Drift Applications*. 91–114 (2016). **11.** Data is from the author’s employer, University of Technology Sydney (UTS). **12.** Human Ethics approval granted by The university of Sydney under protocol# 2017-400

Acknowledgements: feedback from Kalina Yacef & the Computer Human Adapted Interaction (CHAI) research group, University of Sydney;