

**WIH3001**

Data Science Project Proposal

**Semester 1 Session 2023/2024**

**Project Title:**

Online Technical Assessment Classification

Based on Personalization

**Supervisor:**

Associate Prof. Dr. Nor Liyana Bt Mohd Shuib

**Prepared by:**

Lim Zheng Yu

U2102809/1

**TABLE OF CONTENTS**

|  |  |
| --- | --- |
| INTRODUCTION/ BACKGROUND | 3 |
| PROBLEM STATEMENT | 5 |
| OBJECTIVES | 6 |
| DATA SCIENCE METHODOLOGY | 7 |
| REFERENCES | 8 |

**INTRODUCTION/ BACKGROUND**

The unprecedented COVID-19 pandemic has led to a paradigm shift in various aspects of society. It poses as a major health crisis to the world, and it has since hindered society’s progress of development across all areas. With global economies showing signs of failure, the world was stimulated and forced to be more innovative to overcome the various obstacles that the pandemic has brought in. One of the affected fields is the education sector. The delay in education may be detrimental to society, and hence the urgency to continue education work has caused institutions across the globe to bring their teaching into the online realm (Abduh, 2021). This sudden switch to online education means coursework, assignments, as well as assessments have also been transferred to an online platform, with the problem that online education is currently still at its infancy, meaning that course evaluations are mostly uncertain. Work submissions or online tests can be easily cancelled, disabling students to properly assess their capabilities in their studies. These interruptions will eventually cause whole batches of students to perform poorly as compared to other cohorts and will be a huge disadvantage to them when entering the workforce in the future (Sievertsen et al., 2021).

Coincidentally, being in the digital age means that an abundance of new technologies is being developed regularly. The advent of 5G networks and Wi-Fi 6, with its widespread adoption around the time of the pandemic has enabled a smooth transition towards a digitized education. Furthermore, in this era, there is an ever-increasing demand for technical skills in the field of the technological industry. Online technical assessments for technical courses such as in the field of computer science and engineering have since become imperative. Throughout the period of the pandemic, a study by Stamatios Papadakis (2023) shows that there is an unexpected spike in the number of registrations in Massive Open Online Courses (MOOC), which can be attributed to the increased interest in online education during the pandemic. This further proves that there is an immediate necessity for a well-implemented online education system across all platforms to ensure the integrity of quality education.

One problematic issue with assessments is that it can be merely used as a tool to grade and segregate exam candidates into different categories of how well they’ve performed. Thus, the Sun’s Java Certification Examination and Novell’s certification examinations are multiple choice questions (Roberts et al., 2003), serving no significant purpose to evaluate the competency of programming students. It is important to note that technical skills such as programming are more hands-on than theoretical, therefore in the instance of programming assessments, it is wiser to simulate a real-life programming environment, which demonstrates the candidates’ ability to perform programming tasks in real-time, further preparing students for entering the workforce.

Another point to make on traditional assessments is that its standardized format makes them “indirect and inauthentic” (Bailey, 1998). Traditional assessments are effective in measuring students’ capabilities at one point in time but cannot inform educators on how well a student is progressing in their education (Dikli, 2003). Everyone, every student is different in terms of their ability to showcase their skills and absorb information, which can be described through the VAK (Visual-Auditory-Kinesthetic) Learning Styles Model. Some are better in their practical skills; some have better visual memory and hence prefer to learn through observation. This just goes to show that a personalized online technical assessment is much needed to properly evaluate students’ learning, especially in the field of computer science and engineering.

**PROBLEM STATEMENT**

Despite technical assessments being shifted onto online platforms, the root of the issue with them is still not directly addressed. Traditional assessment methods are still being employed, the only difference this time around is the location where they are held. Assessments are typically knowledge-based, which measure a candidate’s information retention capabilities of a particular subject, instead of assessing his ability to perform a task (Maennel, 2019). Imagine a programmer only being able to explain various advanced programming concepts, but unable to type out code segments for tasks. Statistics of the research conducted by Abduh (2021) show that current online assessment methods have a low score in measuring students’ achievements fairly. Thus, showing a need for a personalized online technical assessment for students.

To make matters worse, the teaching of large programming classes means that students have limited interactions with their lecturers and supervisors (Robinson et al., 2017). Minimal guidance in studies especially in courses which demand a steep learning curve from students like programming converts to a high likelihood that students will not be able to fully grasp the essence of classes, which also leads to a high likelihood of failing those classes.

Currently, the most common form of online assessments is the multiple-choice format (Kim et al., 2008), likely due to its easy implementation. In the context of technical assessments, these forms of tests are unsuitable to evaluate one’s capabilities to execute tasks. As mentioned in the VAK model, everyone has a different learning style and hence everyone demonstrates their skills in a different manner. Aside from the fact that multiple-choice format is mostly not suited for technical tests, this just shows that there is an essential requirement for a tailor-made test format for everyone to grade students accurately.

**OBJECTIVES**

There are two objectives in this data science project:

1. To develop a classification model of online technical assessment based on personalization.
2. To evaluate the performance of the classification model of online technical assessment based on personalization.

**DATA SCIENCE METHODOLOGY**

One of the most used data science methodologies is the CRISP-DM methodology, which stands for Cross Industry Standard Process for Data Mining.

1. Business Understanding
   1. Understanding objectives and requirements of project
      1. Set objectives
      2. Produce project plan
   2. What does business needs
2. Data Understanding
   1. Data collection
   2. Describe data
   3. Explore data
   4. Verify data quality
3. Data Preparation
   1. Select data
   2. Clean data
   3. Construct data
      1. Derive new attributes/ data transformation
   4. Integrate data
      1. Create new datasets by combining data from multiple sources
   5. Format data
4. Modelling
   1. Select suitable modelling techniques
      1. Tweak if assessment is unsatisfactory
   2. Generate test design: split data into train test validation sets
   3. Build model
   4. Assess model
      1. Success criteria?
5. Evaluation
   1. Evaluate business success criteria
   2. Review process
   3. Determine next steps
6. Deployment
   1. Plan deployment
   2. Produce final report

**REFERENCES**

Abduh, M. Y. M. (2021). Full-time online assessment during COVID-19 lockdown: EFL teachers’ perceptions. *Asian EFL Journal*, *28*(1.1), 26-46.

Bailey, K. M. (1998). Learning about language assessment: dilemmas, decisionjs, and directions. Heinle& Heinle: US.

Dikli, S. (2003). Assessment at a distance: Traditional vs. alternative assessments. *Turkish Online Journal of Educational Technology-TOJET*, *2*(3), 13-19.

Kim, N., Smith, M. J., & Maeng, K. (2008). Assessment in online distance education: A comparison of three online programs at a university. *Online Journal of Distance Learning Administration*, *11*(1), 1-16.

Maennel, O. M. (2019). Predicting student’s success using technical labs as part of university admission to a cyber security program.

Papadakis, S. (2023). MOOCs 2012-2022: An overview. *Advances in Mobile Learning Educational Research*, *3*, 682-693. https://doi.org/10.25082/AMLER.2023.01.017

Roberts, G. H., & Verbyla, J. L. (2003, January). An online programming assessment tool. In *Proceedings of the fifth Australasian conference on Computing education-Volume 20* (pp. 69-75).

Robinson, P. E., & Carroll, J. (2017, 25-28 April 2017). An online learning platform for teaching, learning, and assessment of programming. 2017 IEEE Global Engineering Education Conference (EDUCON) (pp. 547-556). IEEE. https://10.1109/EDUCON.2017.7942900

Sievertsen, H. H., & Burgess, S. (2020, April 1). *Schools, skills, and learning: The impact of covid-19 on Education*. CEPR. https://cepr.org/voxeu/columns/schools-skills-and-learning-impact-covid-19-education