Ölçme ve Değerlendirme Analitikleri

Okan Bulut

University of Alberta

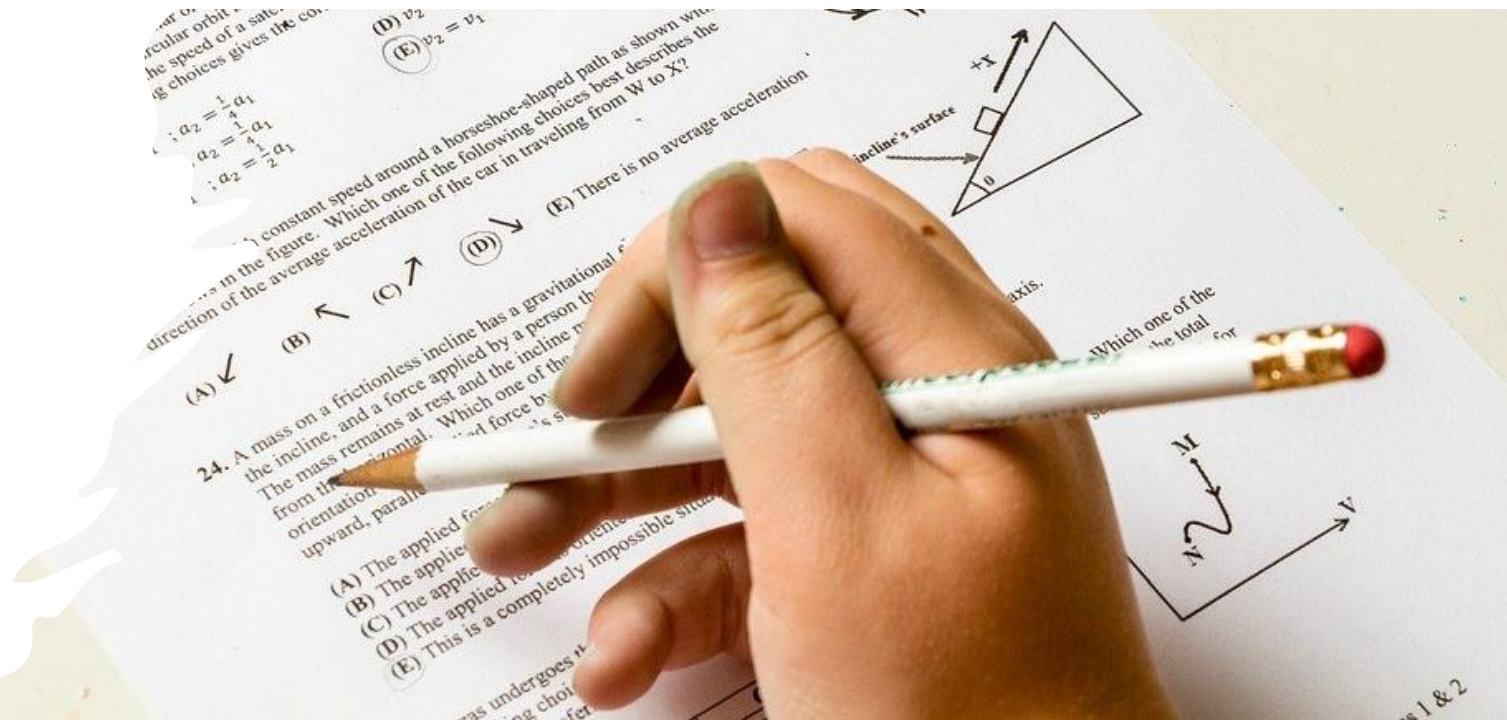✉ bulut@ualberta.ca　　💻 www.okanbulut.com　　🐦 @drokanbulut

Ölçen/Değerlendiren → Süreci/Verileri Yorumlayan

Dijital Ölçme

Bilgisayar destekli, çevrimiçi sınavlar ve ödevler

Soruları cevaplama süreleri

Sorulara verilen cevaplar

Farklı soru türleriyle etkileşim

Çözüm yollarında izlenen stratejiler

Log Verileri
(Dijital Ayak İzleri)

Ödev/sınav bitirme süreleri

Çevrimiçi derslere katılım istatistikleri

Öğrenciler arası etkileşim verileri
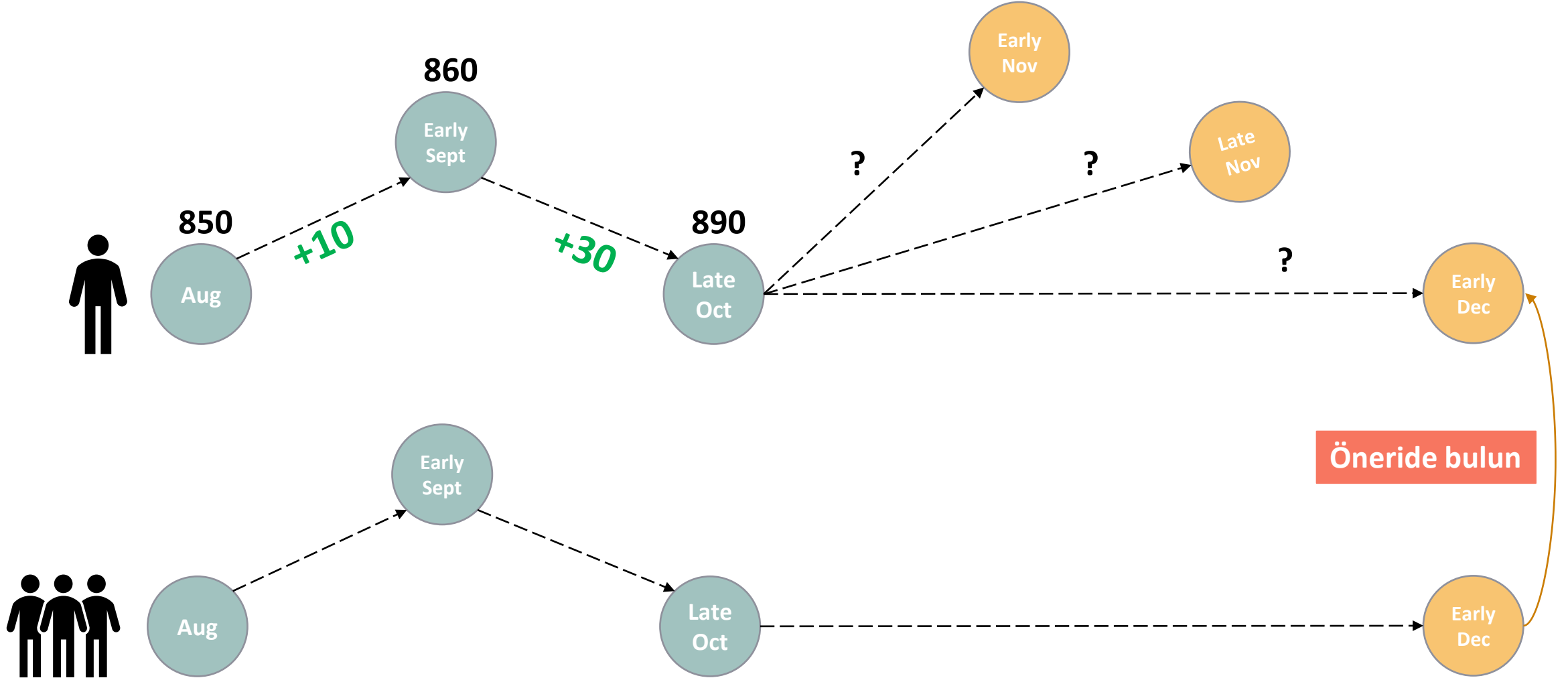
Materyallere erişim sıklığı

Boylamsal veriler

Bireye Özgü Sınav Takvimi Belirleme

Öneri Oluşturma

Paper 1 (left)

An Intelligent Recommender System for Personalized Test Administration Scheduling With Computerized Formative Assessments

Okan Bulut[1]*, Damien C. Cormier[1] and Jinnie Shin[2]

[1] Centre for Research in Applied Measurement and Evaluation, Faculty of Education, University of Alberta, Edmonton, AB, Canada, [2] Department of Educational Psychology, Faculty of Education, University of Alberta, Edmonton, AB, Canada

The introduction of computerized formative assessments in schools has enabled the monitoring of students' progress with more flexible test schedules. Currently, the timing and frequency of computerized formative assessments are determined based on districts and school authorities' agreements with testing organizations, the teachers' judgment of students' progress, and grade-level testing guidelines recommended by researchers. However, these practices often result in a rigid test scheduling that disregards the pace at which students acquire knowledge. Furthermore, students are likely to experience the loss of instructional time due to frequent testing. To administer computerized formative assessments efficiently, teachers should be provided systematic guidance on finding an optimal testing schedule based on each student's progress. In this study, we aim to demonstrate the utility of intelligent recommender systems (IRSs) for generating individualized test schedules for students. Using real data from a large sample of students in grade 2 (n = 355,078) and grade 4 (n = 390,336) who completed the Star Math assessment during the 2017–2018 school year, we developed an IRS and evaluated its performance in finding a balance between data quality and testing frequency. Results indicated that the IRS was able to recommend a fewer number of test administrations for both grade levels, compared with standard practice. Further, the IRS was able to maximize the score difference from one test administration to another by eliminating the test administrations in which students' scores did not change significantly. Implications for generating personalized schedules to monitor student progress and recommendations for future research are discussed.

Keywords: recommender system, formative assessment, personalized learning, progress monitoring, mathematics

INTRODUCTION

Classroom assessments allow K–12 teachers to evaluate student learning (i.e., monitor students' progress) and make a variety of important decisions about learning outcomes (e.g., producing feedback, assigning grades). Teachers use two types of assessments to evaluate student learning in the classroom: summative and formative (Black and Wiliam, 2009). Summative assessments are typically used *at the end* of an instructional period, such as a unit or a semester, to evaluate student learning relative to content standards or benchmarks. Quizzes, midterm exams, or a final project are

https://www.frontiersin.org/articles/10.3389/feduc.2020.572612/full

---

Paper 2 (right)

Building an intelligent recommendation system for personalized test scheduling in computerized assessments: A reinforcement learning approach

Jinnie Shin[1] · Okan Bulut[2]

Abstract

The introduction of computerized formative assessments in the classroom has opened a new area of effective progress monitoring with more accessible test administrations. With computerized formative assessments, all students could be tested at the same time and with the same number of test administrations within a school year. Alternatively, the decision for the number and frequency of such tests could be made by teachers based on their observations and personal judgments about students. However, this often results in rigid test scheduling that fails to take into account the pace at which students acquire knowledge. To administer computerized formative assessments efficiently, teachers should be provided with systematic guidance regarding effective test scheduling based on each student's level of progress. In this study, we introduce an intelligent recommendation system that can gauge the optimal number and timing of testing for each student. We discuss how to build an intelligent recommendation system using a reinforcement learning approach. Then, we present a case study with a large sample of students' test results in a computerized formative assessment. We show that the intelligent recommendation system can significantly reduce the number of testing for the students by eliminating unnecessary test administrations where students do not show significant progress (i.e., growth). Also, the proposed recommendation system is capable of identifying the optimal test time for students to demonstrate adequate progress from one test administration to another. Implications for future research on personalized assessment scheduling are discussed.

Keywords Test administration optimization · Reinforcement learning · Computerized formative assessment · Personalized learning

Introduction

Computerized formative assessments establish a connection between computer technologies and formative assessments to effectively manage and deliver classroom assessments (Webb et al., 2013). Formative assessments refer to the general process that engages students and teachers to evaluate student learning, provide feedback, and improve learning outcomes using formative evaluation (McManus, 2008). In a critical review of research on formative assessments, Dunn and Mulvenon (2009) collectively defined formative evaluation as the evidence-based evaluation aiming to "inform teachers, students, and educational stakeholders about the teaching and learning process" (p.4, Dunn & Mulvenon, 2009). A diverse format of tools can be applied to formative assessment to effectively collect evidence of student learning (Bennett, 2011). Such tools include, but are not limited to, peer feedback and assessment (e.g., Volante & Beckett 2011), questioning (e.g., Black & Harrison, 2001), self-assessment (e.g., Andrade, 2019), and pedagogical documentation, such as portfolios (e.g., Buldu, 2010).

Appropriate use of formative assessment tools in the classroom has previously been identified to result in a profound student learning (McManus, 2008; Kingston & Nash, 2011). Meta-analyses on the effect of classroom formative assessment have revealed that formative assessments show predominant effects to improve learning outcomes

✉ Jinnie Shin
jinnie.shin@coe.ufl.edu

Okan Bulut
bulut@ualberta.ca

[1] Institute for Advanced learning Technologies, College of Education, University of Florida, Gainesville, FL, USA

[2] Centre for Research in Applied Measurement and Evaluation, University of Alberta, Alberta, Canada

Springer

https://link.springer.com/article/10.3758/s13428-021-01602-9

Dinlediğiniz için
teşekkür ederim.