USER BEHAVIOR MODELING

Ashay Tamhane
Staff Data Scientist, Swiggy

USER MODELS?

REPRESENTATION OF A USER

USER MODELS?

WHY MODEL USERS?

WHY MODEL USERS?

If we can represent what users did historically as a probabilistic model,

we can predict what they might do in future

If we can **predict** what they might do in future,

we can take appropriate **actions**

WHY IS IT A CHALLENGE?

WHY IS IT A CHALLENGE?

Predicting online user behavior is like solving a mystery

Guess the behavior with extremely scarce information from their digital footprint

TALK STRUCTURE

Food: Why does user behavior matter?

Fashion: How to incorporate user behavior in approach?

Education: Predicting user behavior in non-retail domain

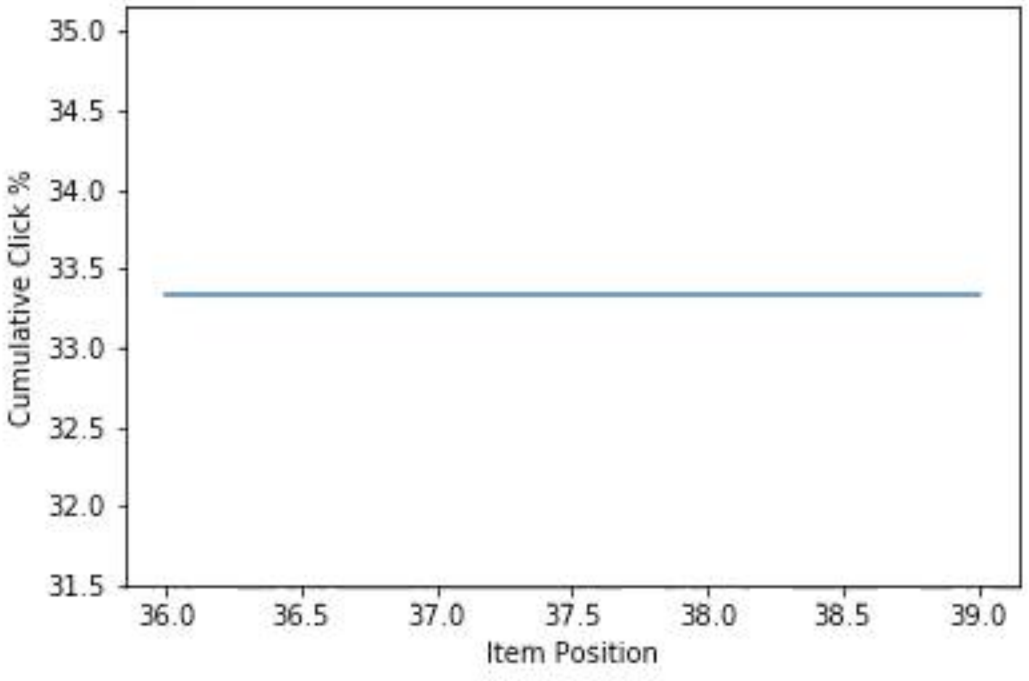
Food

UNDERSTANDING USER PATTERNS

QUESTION: Do Positions of Items on Menu Bias User Decisions?

UNDERSTANDING USER PATTERNS

QUESTION: Do Positions of Items on Menu Bias User Decisions?

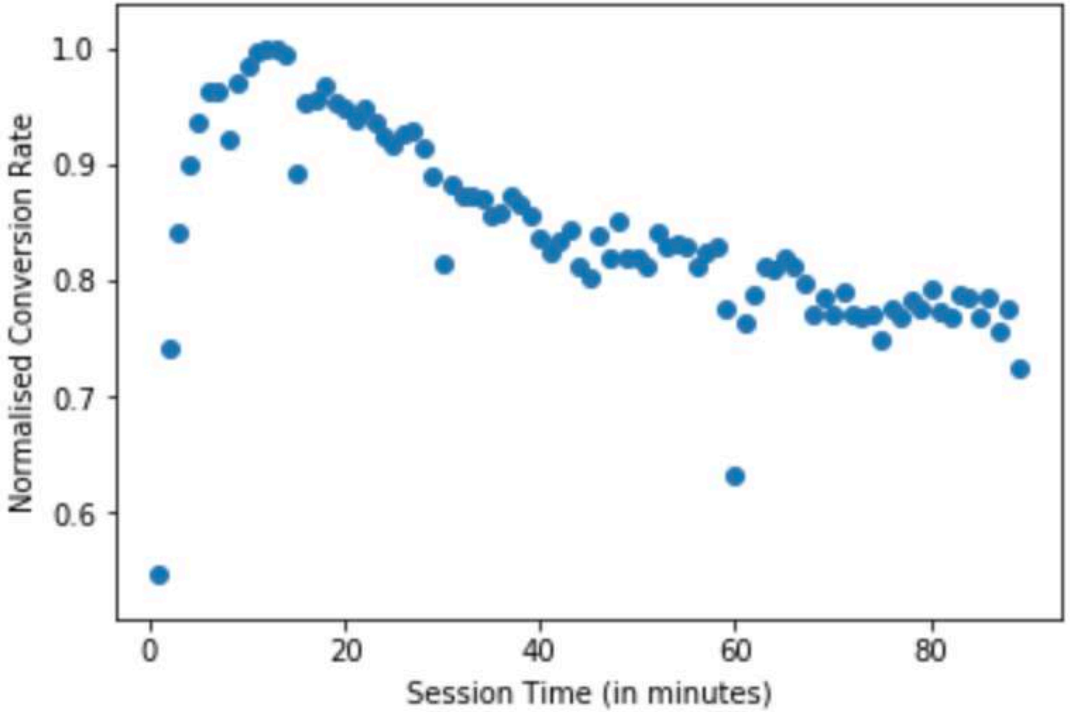


UNDERSTANDING USER PATTERNS

QUESTION: IS IT GOOD IF USERS SPEND MORE TIME ON PLATFORM?

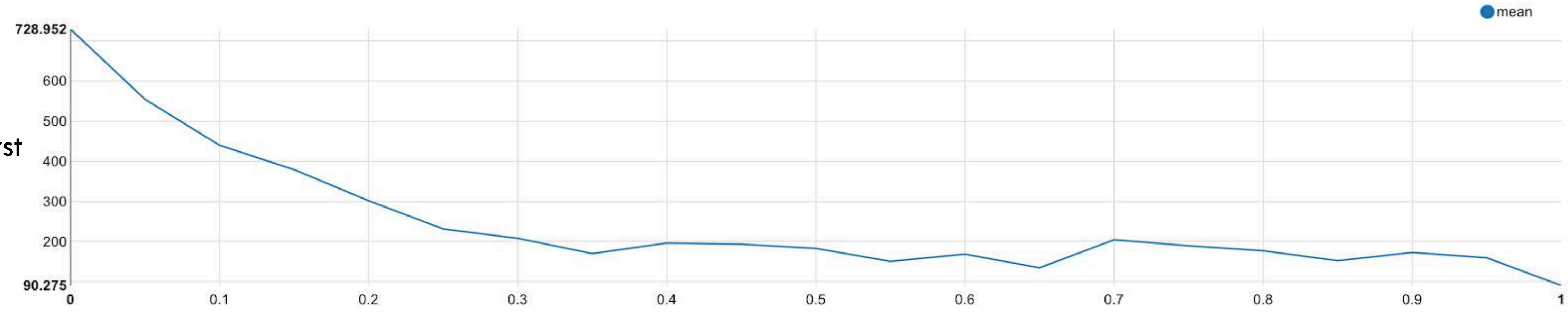
UNDERSTANDING USER PATTERNS

QUESTION: IS IT GOOD IF USERS SPEND MORE TIME ON PLATFORM?



UNDERSTANDING USER PATTERNS

Time for first click



NDCG

NDCG to Conversion graph flips

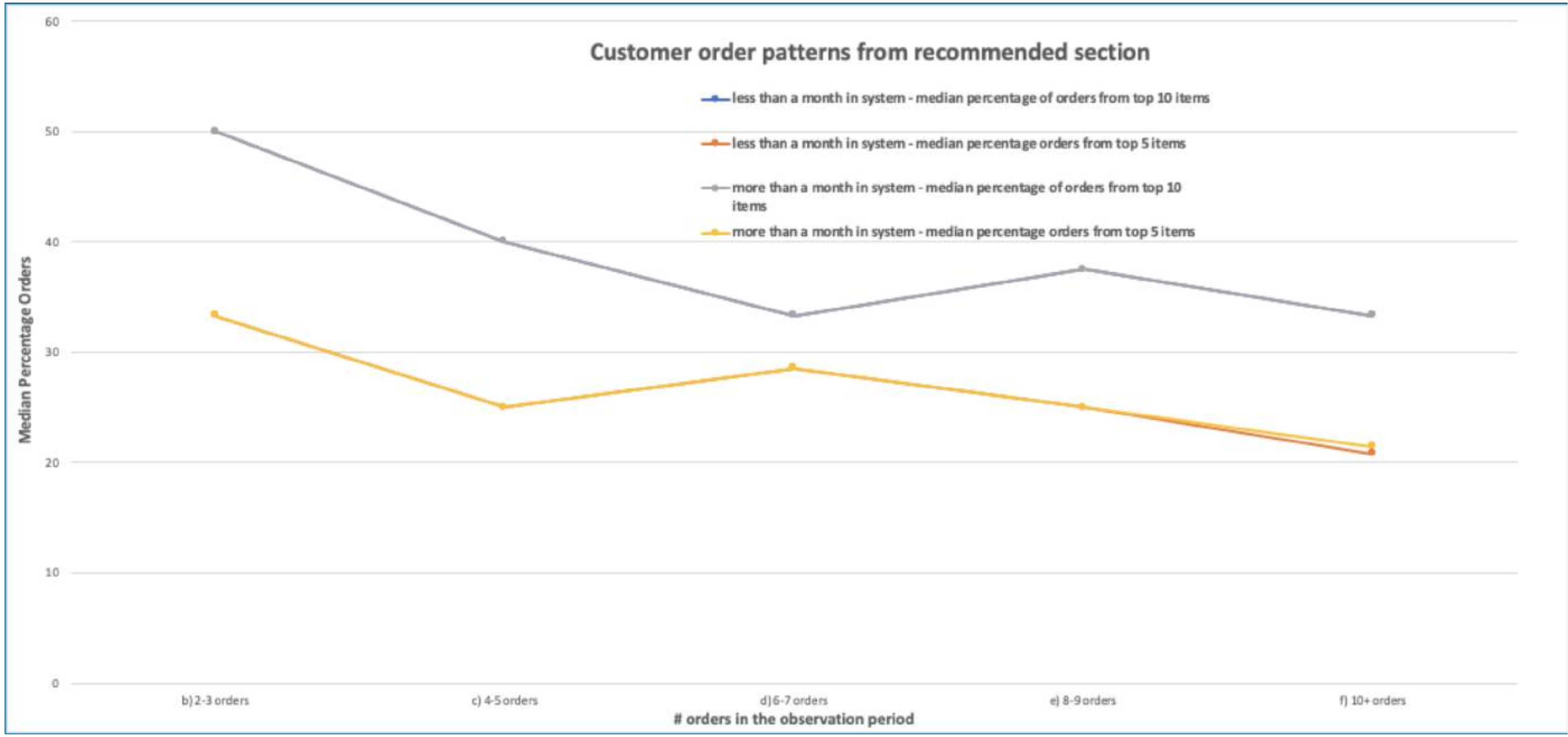
UNDERSTANDING USER PATTERNS

QUESTION: Do users order only popular items?

UNDERSTANDING USER PATTERNS

QUESTION: Do users order only popular items?

Orders from Top 5 Items	Orders from Top 10 Items
28%	42%



IMPLICATIONS ON RECOMMENDATIONS

- Most users know what they want to eat
- Need to predict intent and show relevant ranking to avoid drop offs
- Different user segments need to be handled differently

	New User on Platform	Old User on Platform
New Restaurant on Platform	No Historical Context for both restaurant or user **Case 1**	Have historical user context for some items, which may or may not map to new restaurant. **Case 2**
Old Restaurant on Platform, But New Restaurant for User	Have historical context for restaurant, but no context for user **Case 3**	Have historical user context, which may or may not map to new restaurant **Case 4**
Old Restaurant on Platform, Old Restaurant for User	**NA**	Have historical context for both restaurant and user **Case 5**

ALGORITHMS

- Baseline: Popularity
- Collaborative Filtering
- Learning to Rank: pointwise, pairwise, listwise
- Multi arm bandits
  - Non Contextual
  - Contextual
  - Full Reinforcement Learning

Fashion

IMPLICATIONS ON RECOMMENDATIONS

- ~~Most users know what they want to buy~~
- Need to predict intent and show relevant ranking to avoid drop offs
- Different user segments need to be handled differently
- Bigger sessions are better

Modeling Contextual Changes in User Behaviour in Fashion e-Commerce

Ashay Tamhane(✉), Sagar Arora, and Deepak Warrier

Myntra, Bengaluru, India
{ashay.tamhane,sagar.arora,deepak.warrier}@myntra.com

Abstract. Impulse purchases are quite frequent in fashion e-commerce; browse patterns indicate fluid context changes across diverse product types probably due to the lack of a well-defined need at the consumer's end. Data from our fashion e-commerce portal indicate that the final product a person ends-up purchasing is often very different from the initial product he/she started the session with. We refer to this characteristic as a 'context change'. This feature of fashion e-commerce makes understanding and predicting user behaviour quite challenging. Our work attempts to model this characteristic so as to both detect and preempt context changes. Our approach employs a deep Gated Recurrent Unit (GRU) over clickstream data. We show that this model captures context changes better than other non-sequential baseline models.

1 Introduction

Understanding user behaviour is critical for any e-commerce platform in order to personalise products and induce the user to convert. This becomes easier if the user has a well-defined need and exhibits cohesive intent. Unlike other domains, purchases are often impulsive in fashion e-commerce. While some users do visit a fashion portal out of their need to purchase specific products, a huge number of folks are there just to explore and transact once they come across a product of their liking. This results in "fluid" browsing patterns that cut across different products categories. In other words, there are many context changes that happen in a typical user session. This makes modeling the user's behaviour, and hence personalisation complex [16, 26].

Consider real user sessions illustrated in Figs. 1 and 2. In the session shown in Fig. 1, the user browsed a Shirt, and on the very next click switched to Shorts. Again on the third click, we see a switch back to a Shirt (of a different style altogether) and then again a switch back to a few Shorts in the next clicks. After consistently viewing a few Shorts, the user switches to a T-shirt on his final click. Such changing contexts makes modeling the user behaviour quite difficult and simple heuristics will fall short. For example, if we predicted that the user would browse Shorts in the final click using majority voting, it would have failed. It is important to note that before browsing Shorts, the user also looked at a couple of Shirts and later changed context to Shorts. This is indicative of the fact that the user wasn't really looking for Shorts from the moment he started browsing

J. Kim et al. (Eds.): PAKDD 2017, Part II, LNAI 10235, pp. 539–550, 2017.
DOI: 10.1007/978-3-319-57529-2_42

BACKGROUND

- India’s leading fashion e-commerce portal
  - Apparel, Footwear, Accessories
  - Tshirts, Shirts, Jeans, Dresses, etc (**“product categories”)**

- Product catalogue is large and dynamic
  - **408,155** products for Men & **625,171** products for Women
  - Around **2000** products are added every day

To help users find the most relevant products,

Personalization Is Critical

SIMILARITY BASED APPROACH








65% of the daily 1.3 million sessions have more than one product category




In 41% of the sessions the purchased product is different from first browsed product

PROBLEM DEFINITION

Lack of a well defined need in fashion e-commerce

User behaviour is impulsive, hence harder to predict

Given a user session $\boldsymbol{p_1, p_2, \ldots p_{n-1}}$, predict $\boldsymbol{p_n}$

CHALLENGES

- Large and dynamic catalogue
- New styles have no historical data
- “Fluid” context switches are common

Use a combination of *Product Groups* & Sequential Modeling

PRODUCT GROUPS

Blackberrys

2195

Blackberrys

1795

Proline

1449

Proline

1449

Blackberrys

2395

Created from user sessions using word embeddings (skip-gram)

Each Session is a "document" and every word is a product attribute

PRODUCT GROUPS

Aamii

1499

Aamii

1499

RANGMANCH BY PANTALOONS

599

RANGMANCH BY PANTALOONS

599

PROBLEM DEFINITION

Lack of a well defined need in fashion e-commerce

User behaviour is impulsive, hence harder to predict

Given a user session $p_1, p_2, \ldots p_{n-1}$, predict p_n

↓

Given a user session $pg_1, pg_2, \ldots pg_{n-1}$, predict pg_n

BASELINES

Majority Voting

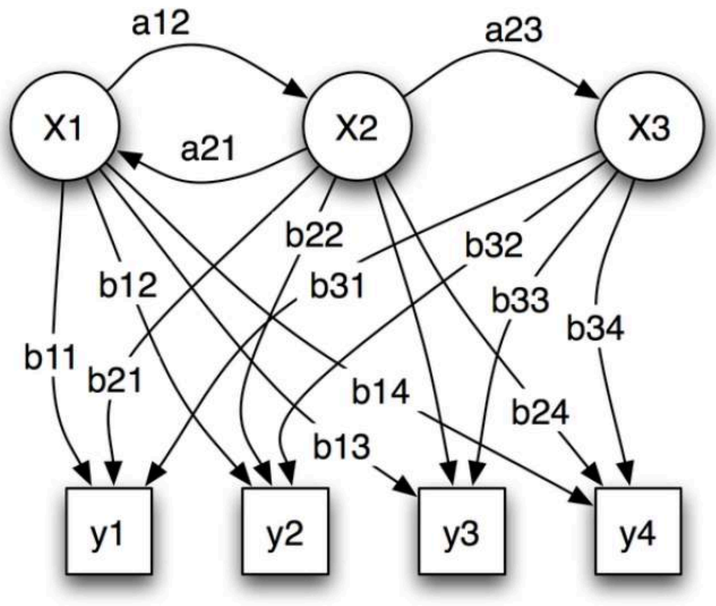
- Replace every product by it's product group (similarity based approach)
- Select product groups in the order of their frequencies

Product Group Graph (PG Graph)

- Create a directional graph (nodes are product groups, edges are normalized transition counts)
- Takes into account transitions between product groups (Men Tshirts -> Men Shorts)

Doesn't take sequential context into consideration

SEQUENTIAL MODELING WITH HMM



Source: Rabiner

X1:

Shorts: 0.4

Sunglasses: 0.3

Hat: 0.2

Others: 0.2

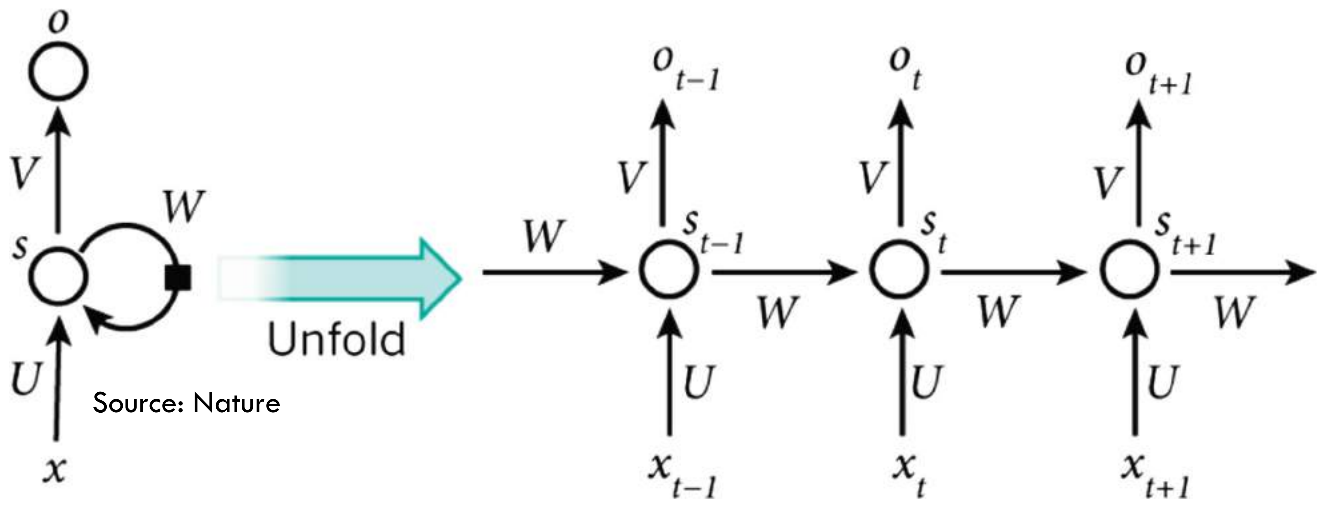
X1: Summer/Beach State

Sample Training Sequence




SEQUENTIAL MODELING WITH RNN

Output at step t (probability vector across vocab)



Source: Nature

One hot encoded

Suffers from vanishing gradient problem

↓

LSTM, **GRU**

MODELING CONTEXTUAL CHANGES WITH GRU

$pg_1, pg_2, \ldots pg_{n-1} \longrightarrow \boxed{\text{GRU}} \longrightarrow pg_i, pg_{i+1}, \ldots, pg_n$

Given the sequence of product groups, get the probability of the next product group

RESULTS

Table 1. MRR and NDCG as a function of K for Men

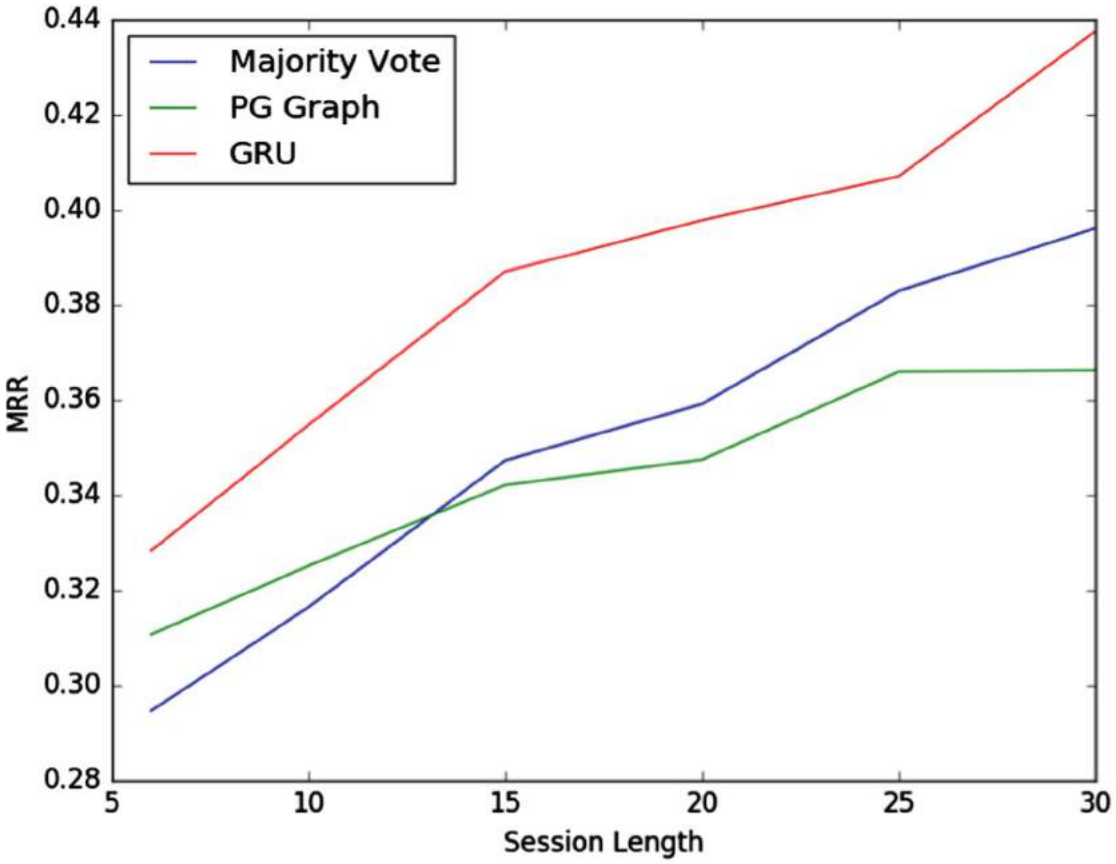
	GRU		PG graph		Majority voting	
	MRR	NDCG	MRR	NDCG	MRR	NDCG
K = 3	$0.32_{(+15.16\%,+10.5\%)}$	$0.20_{(+26\%,+1.9\%)}$	0.28	0.16	0.29	0.20
K = 5	$0.34_{(+12.17\%,+8.7\%)}$	$0.18_{(+17.3\%,+1.2\%)}$	0.30	0.16	0.31	0.18
K = 10	$0.35_{(+10.34\%,+10.51\%)}$	$0.19_{(+13.87\%,+44.49\%)}$	0.31	0.17	0.31	0.13

Table 2. MRR and NDCG as a function of K for Women

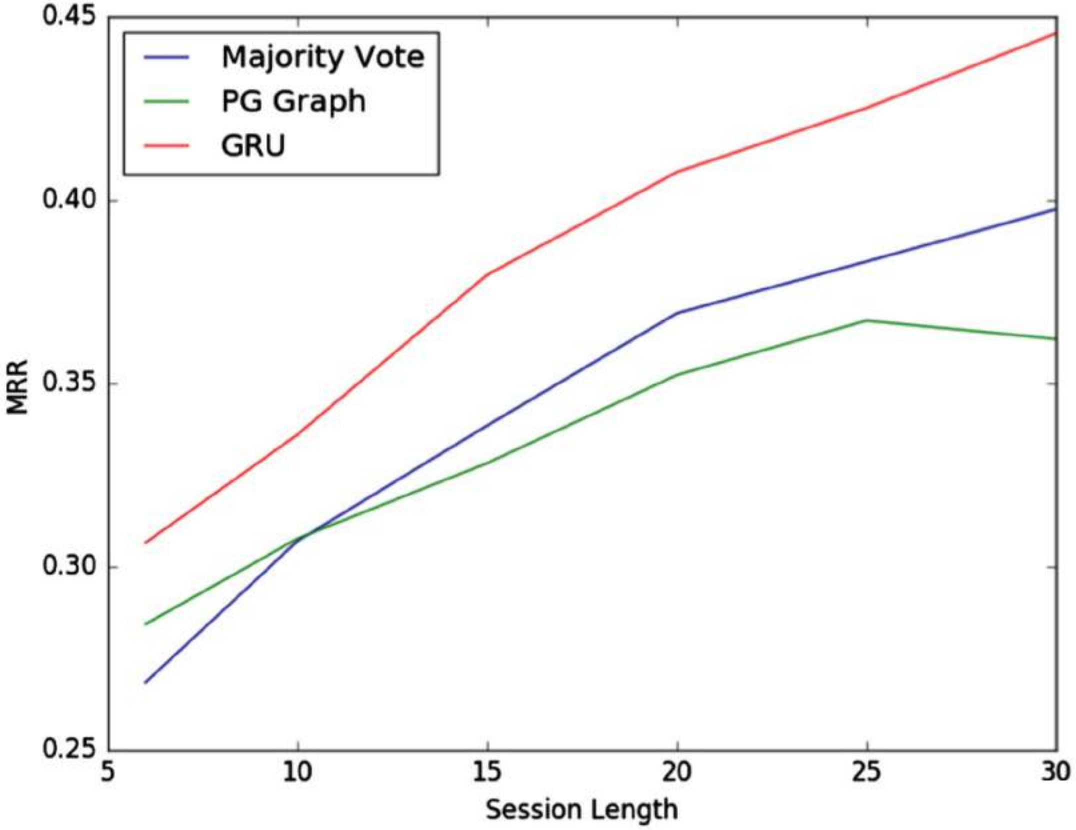
	GRU		PG graph		Majority voting	
	MRR	NDCG	MRR	NDCG	MRR	NDCG
K = 3	$0.31_{(+17.94\%,+9.5\%)}$	$0.28_{(+0.99\%,+304\%)}$	0.26	0.28	0.28	0.07
K = 5	$0.32_{(+15\%,+7.6\%)}$	$0.28_{(+0.3\%,+123\%)}$	0.28	0.28	0.30	0.12
K = 10	$0.34_{(+12.72\%,+9.5\%)}$	$0.26_{(+13.64\%,+144\%)}$	0.30	0.23	0.31	0.11

MRR : Mean Reciprocal Rank
NDGC : Normalised Discounted Cumulative Gain
K : Number of recommendations

RESULTS



Men

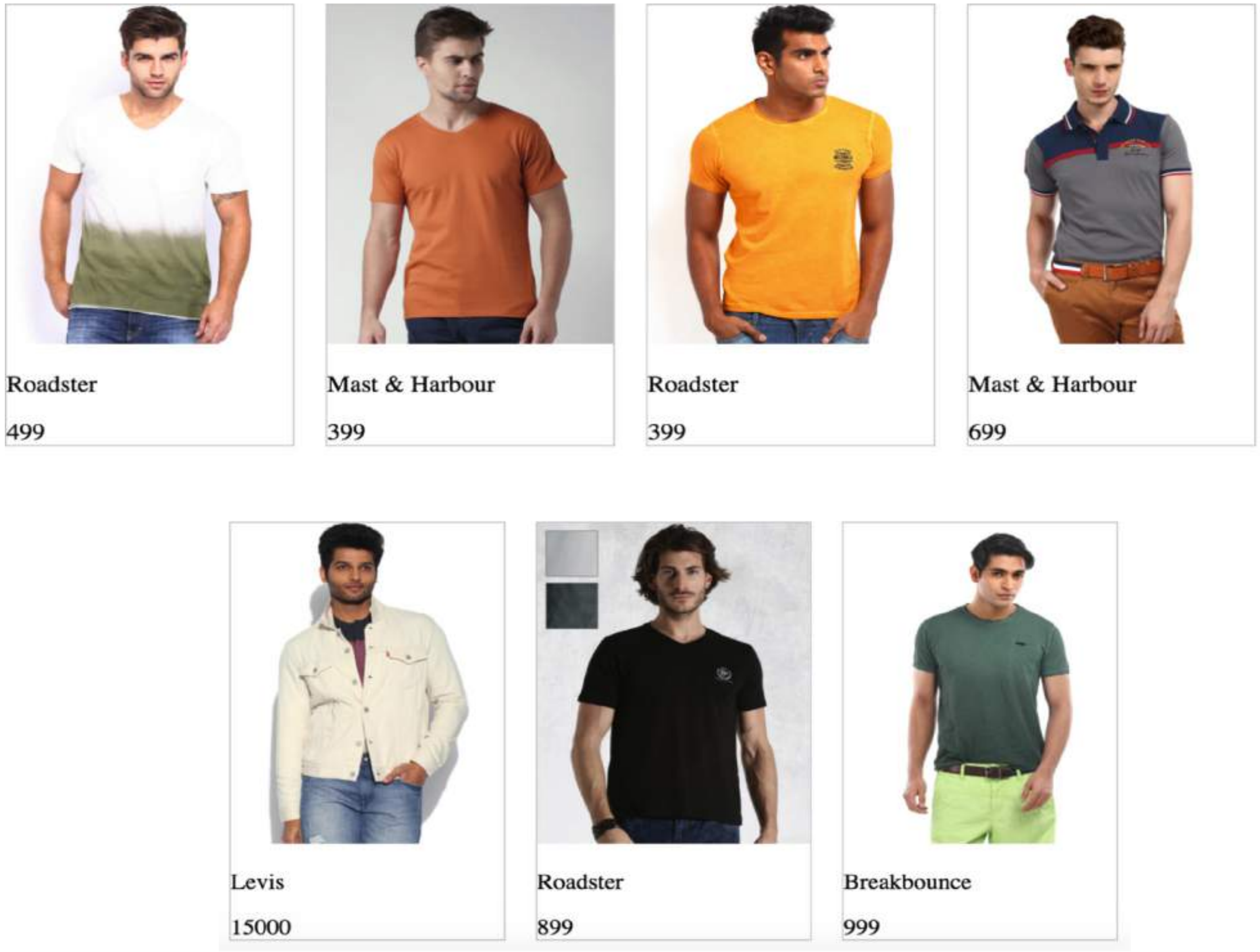


Women

MODELING CONTEXTUAL CHANGES WITH GRU

Artificially Generated Sample from GRU

MODELING CONTEXTUAL CHANGES WITH GRU



Artificially Generated Sample from GRU

SUMMARY

- Context changes are common in fashion e-commerce
- In such a scenario, sequential modeling with GRU performs better than non-sequential baselines across different session lengths

Education

Predicting Student Risks Through Longitudinal Analysis

Ashay Tamhane
IBM Research - India
ashayuda@buffalo.edu

Shajith Ikbal
IBM Research - India
shajmoha@in.ibm.com

Bikram Sengupta
IBM Research - India
bsengupt@in.ibm.com

Mayuri Duggirala*
Tata Research Development & Design Centre, Pune, India
mayuri.duggirala@tcs.com

James Appleton
Gwinnett County Public Schools, GA, USA
jim_appleton@gwinnett.k12.ga.us

ABSTRACT

Poor academic performance in K-12 is often a precursor to unsatisfactory educational outcomes such as dropout, which are associated with significant personal and social costs. Hence, it is important to be able to predict students at risk of poor performance, so that the right personalized intervention plans can be initiated. In this paper, we report on a large-scale study to identify students at risk of not meeting acceptable levels of performance in one state-level and one national standardized assessment in Grade 8 of a major US school district. An important highlight of our study is its scale - both in terms of the number of students included, the number of years and the number of features, which provide a very solid grounding to the research. We report on our experience with handling the scale and complexity of data, and on the relative performance of various machine learning techniques we used for building predictive models. Our results demonstrate that it is possible to predict students at-risk of poor assessment performance with a high degree of accuracy, and to do so well in advance. These insights can be used to pro-actively initiate personalized intervention programs and improve the chances of student success.

Keywords

education; educational data mining; risk prediction; longitudinal data analysis

1. INTRODUCTION

One of the primary goals of any education system is to equip students with the knowledge and skills needed to transition to successful career pathways. How effectively education systems around the world are able to meet this goal acts as a major determinant of economic and social progress. In particular, K-12 reflects the most critical phase of an individual's lifelong learning, during which the opportunities for a successful future need to be created and nurtured. It is in recognition of this fact that increasingly, educational reforms focused around frequent and standardized testing of K-12 students are gaining acceptance as a way to monitor performance and progress at an individual and institutional level [3, 4]. For many of these assessments, a passing score is required to meet state or district standards for progression to the next grade level. Poor academic performance in such standardized assessments can thus lead to unfavorable educational outcomes such as grade retention, and when not adequately addressed, it can eventually trigger dropout or sub-optimal career pathways, which are associated with significant personal and social costs. For example, in the United States, nearly 7,000 high school students drop out of school each day; if the students who dropped out of the class of 2011 had graduated, the nation's economy would have benefited from $154 billion in additional income over the course of their lifetimes [11]. Hence, it is important to identify students at risk of poor performance in major standardized tests, and to do so well in advance, so that the right personalized intervention plans can be initiated to improve performance.

Traditionally, K-12 educators (e.g. class teachers) have relied on recent academic results of a student (e.g. in formative tests in the current grade) along with an educator's general intuition gleaned from teaching similar students in the past, to determine if the student might be at risk of poor performance in an upcoming assessment. This makes the process overly reliant on an educator's experience level, there is no objective quantification of the level of risk, and the dependence on recent data to make a prediction for the academic year may often not leave enough time to apply the right level of intervention to adequately improve performance. However, with the digitization of school records and rapid uptake of digital tools for teaching and learning, various aspects of a student's longitudinal journey through K-12 are now captured and persisted in digital form. This offers a rich repository of data that can be analyzed to detect patterns associated with unsatisfactory educational outcomes, derived from thousands of students who have progressed through the system over the years, and taking into account a holistic view of a student in terms of both academic history over a period of time, as well as other non-academic attributes (e.g. related to attendance, demographics, behavior etc.) that may influence academic performance.

*This work was done while the author was at IBM Research-India.

KDD'14, August 24–27, 2014, New York, NY, USA.

http://dx.doi.org/10.1145/2623330.2623355.

Problem & Motivation

- Education domain is witnessing unprecedented transformation
- K-12 schooling – crucial period in everyone's education life
- One of the major problems at K-12 level – drop-outs
- Poor academic performance – One of the key indicators of drop-out
- Predict potential risks in academic performance for early intervention

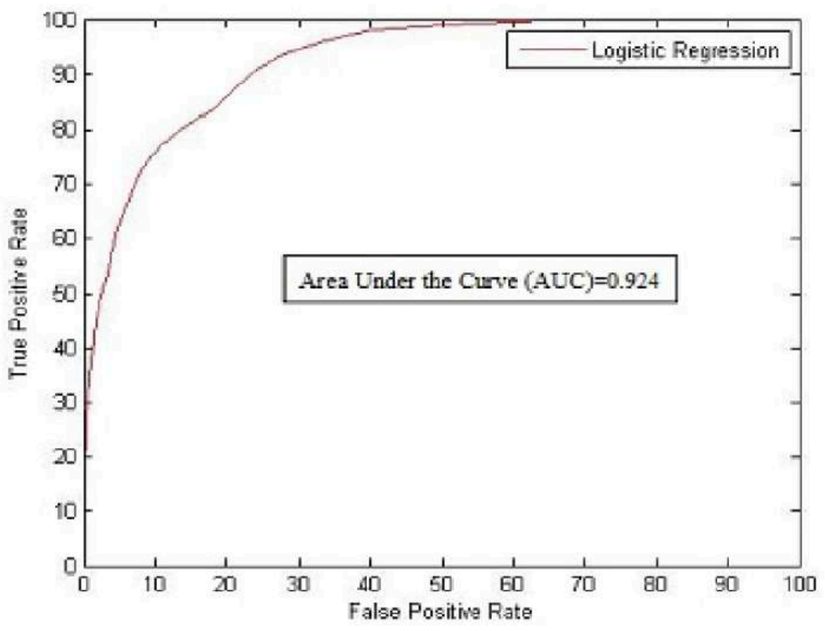
Predicting potential risks in performance of the students ahead in time!

Prediction Task

- Targets considered:
  - CRCT 8th Grade Mathematics
  - CRCT 8th Grade Science
  - ITBS 8th Grade Mathematics
- Data Preparation:
  - Target: for CRCT score < 800 is considered 'at-risk'. For ITBS score < 25 is 'at-risk'
  - Features: all scores from grades < 8th grade + demography + behavior – many scores missing
  - Students chosen such that at least 20% features are present
  - Missing features are mean imputed
  - Data size: CRCT - 58707 students and 342 features; ITBS - 43310 students and 282 features
  - Experimental setup: 5-fold cross validation
- Prediction:
  - Classifiers from IBM SPSS or WEKA: logistic regression, naïve bayes, decision tree
  - To predict: 'at-risk' and 'no-risk' students.
- Evaluation metric:
  - ROC-AUC – area under receiver operating curve - true positives vs false positive

Risk Prediction Performance

Sample ROC curve →



ROC-AUC for various classifiers

↓

Classifier	CRCT 8th Grade Mathematics	CRCT 8th Grade Science	ITBS 8th Grade Mathematics
Naive Bayes	0.744	0.739	0.702
Decision Tree	0.822	0.774	0.766
Decision Table	0.933	0.902	0.893
Logistic Regression	0.924	0.907	0.896

FP for TP>=90 →

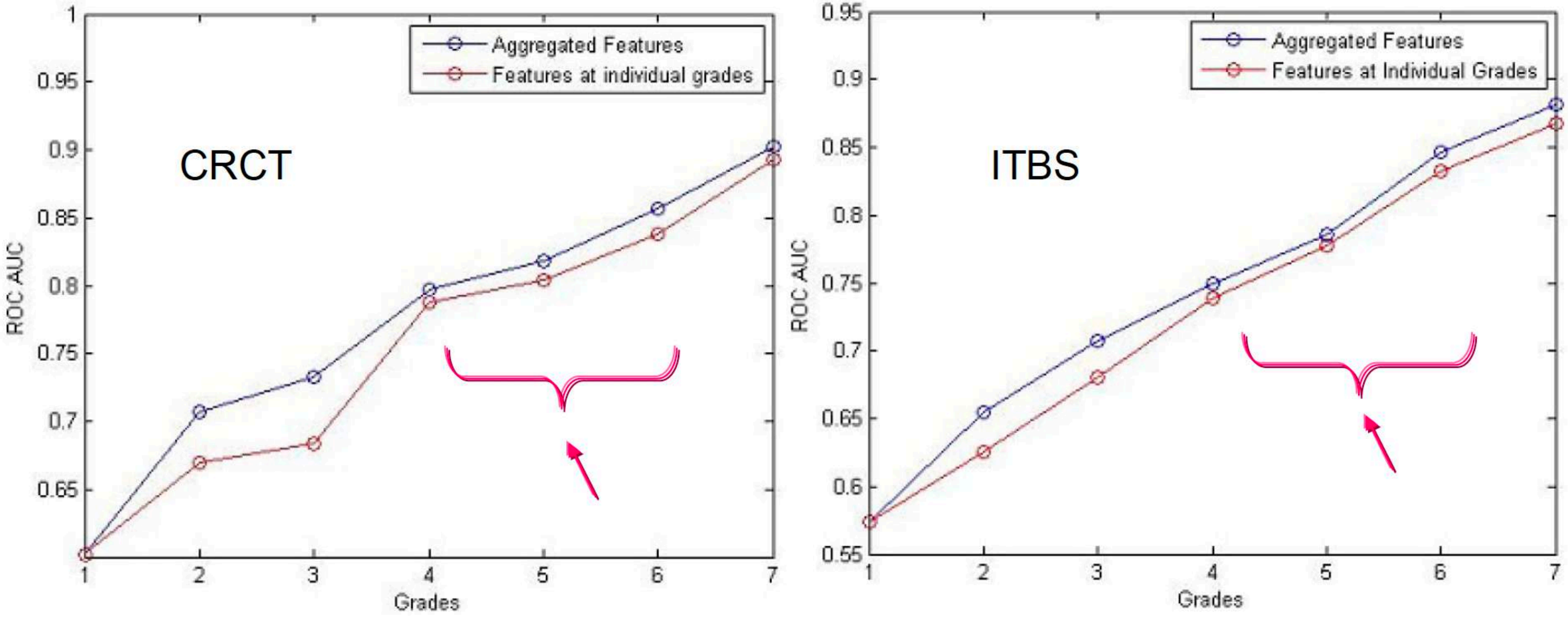
Task	Probability Threshold	True Positive TP, in %	False Positive FP, in %
CRCT 8th Mathematics	0.06	90.5	23.8
CRCT 8th Science	0.18	90.0	24.7
ITBS 8th Mathematics	0.1	90.7	28.8

Feature Importance

Feature Type	CRCT 8th Grade Mathematics	ITBS 8th Grade Mathematics
All Features	0.924	0.896
All Scores	0.902	0.882
All Demographics	0.866	0.814
All Behavioral	0.576	0.559
Scores - Maths	-	0.871
Scores - Science	-	0.828
Scores - Language	-	0.846
Scores - Others	-	0.829
Demography - Gender	0.547	0.537
Demography - Ethnicity	0.660	0.668
Demography - Gifted	0.622	0.630
Demography - Free Meal	0.646	0.640
Demography - Special Education Needs	0.721	0.637
Behavioral - Absence	0.537	0.542
Behavioral - Suspensions	0.588	0.578
Behavioral - Incidents Reported	0.583	0.569

- **Scores are important, demography information helps**
- **Recent past scores are the most important**

Early Prediction



- At Grade 4, it is possible to predict for Grade 8 with reasonably high accuracy
- Accuracy improves as more and more features are aggregated from lower grades