Lab 1 – AskMissy Description

Paul Rankin

Old Dominion University

Computer Science

CS411W

Janet Brunelle

31 January 2022

Version 1

Table of Contents

List of Figures

List of Tables

Introduction

Students' interest in reading is in steady decline. This lack of interest in reading is significantly affecting students' performance in school. The lack of interest in reading is likely caused by students not having access to or not knowing about books that really connect with their interests. Students need a way to find books that can engage with and get them interested in learning. Currently, students are limited in their opportunities to find such books. The most likely way a student will find a book they connect with is through their school, either by teacher or librarian recommendation. This is an inefficient method to search for books of interest. Software can help fill this gap in student learning. This can be achieved through using machine learning to identify student interest similar to how Netflix and other such streaming services recommend shows. As students begin to read more they will inevitably become more interested in learning in addition to the many other benefits of reading. AskMissy is the software we plan to develop to bridge the gap between students and reading. It will utilize machine learning to connect students with books that connect with them.

Societal Problem

When a younger student beginning their academic journey goes to find a book to read they are most likely to turn to teachers, librarians, and their peers. These kinds of people, depending on how well they know the student and how many books they know about, could make a recommendation that works for the student and get them interested in reading overall. However, because this is entirely dependent on both knowing the student well and having a wide array of books to recommend. This method is inefficient and a bad recommendation could lead to a student becoming disheartened with reading in general.

Specifically younger students would likely not know alot about the wide range of book topics and genres. This lack of knowledge closes them off from a whole other world of learning they otherwise could possibly enjoy. Exposing students to a wide variety of topics and genres is important for the learning process at all ages. So a lack of knowledge of what is out there could be detrimental to a student's education.

Many students, specifically those in poorer school systems, might not have access to many books or resources at their local library. This kind of lack of access greatly limits learning to those who may benefit from it the most.
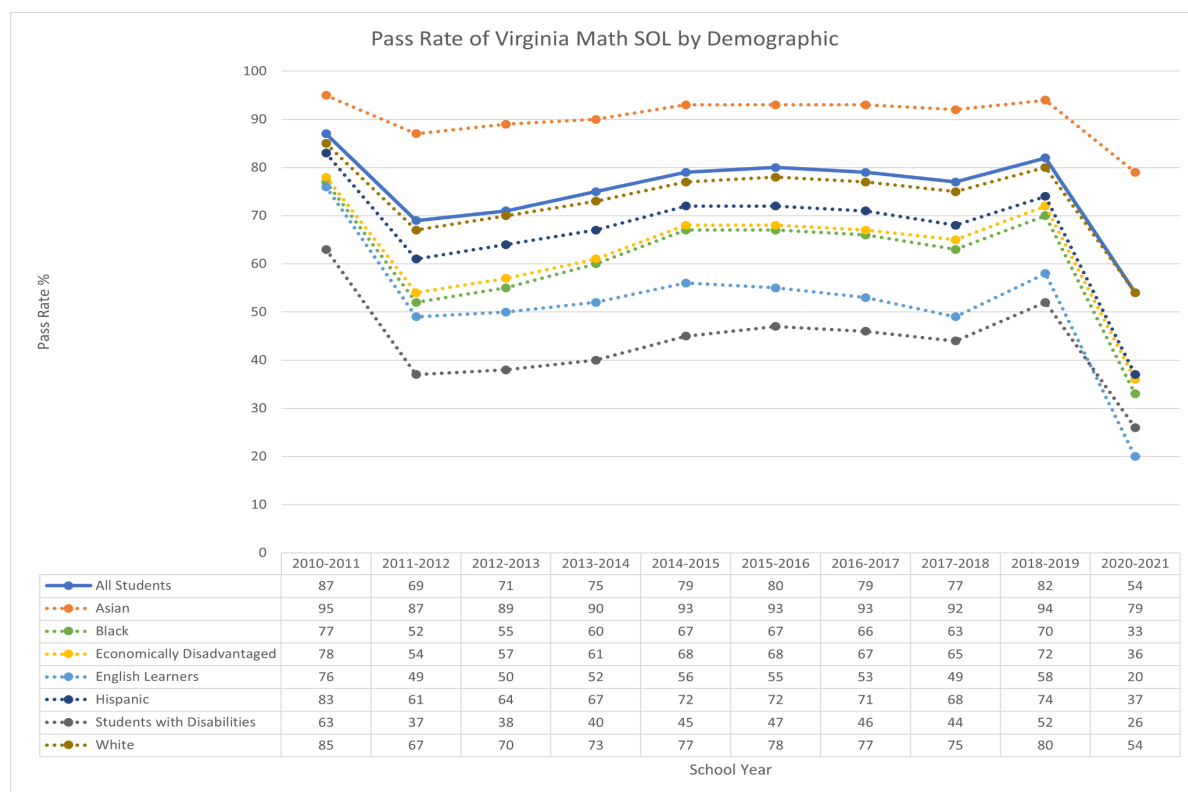
Figure 1:

Pass Rate of Virginia Math SOL by Demographic

	2010-2011	2011-2012	2012-2013	2013-2014	2014-2015	2015-2016	2016-2017	2017-2018	2018-2019	2020-2021
All Students	87	69	71	75	79	80	79	77	82	54
Asian	95	87	89	90	93	93	93	92	94	79
Black	77	52	55	60	67	67	66	63	70	33
Economically Disadvantaged	78	54	57	61	68	68	67	65	72	36
English Learners	76	49	50	52	56	55	53	49	58	20
Hispanic	83	61	64	67	72	72	71	68	74	37
Students with Disabilities	63	37	38	40	45	47	46	44	52	26
White	85	67	70	73	77	78	77	75	80	54

School Year

Figure 2:

Pass Rate of Virginia Reading SOL Tests by Demographic

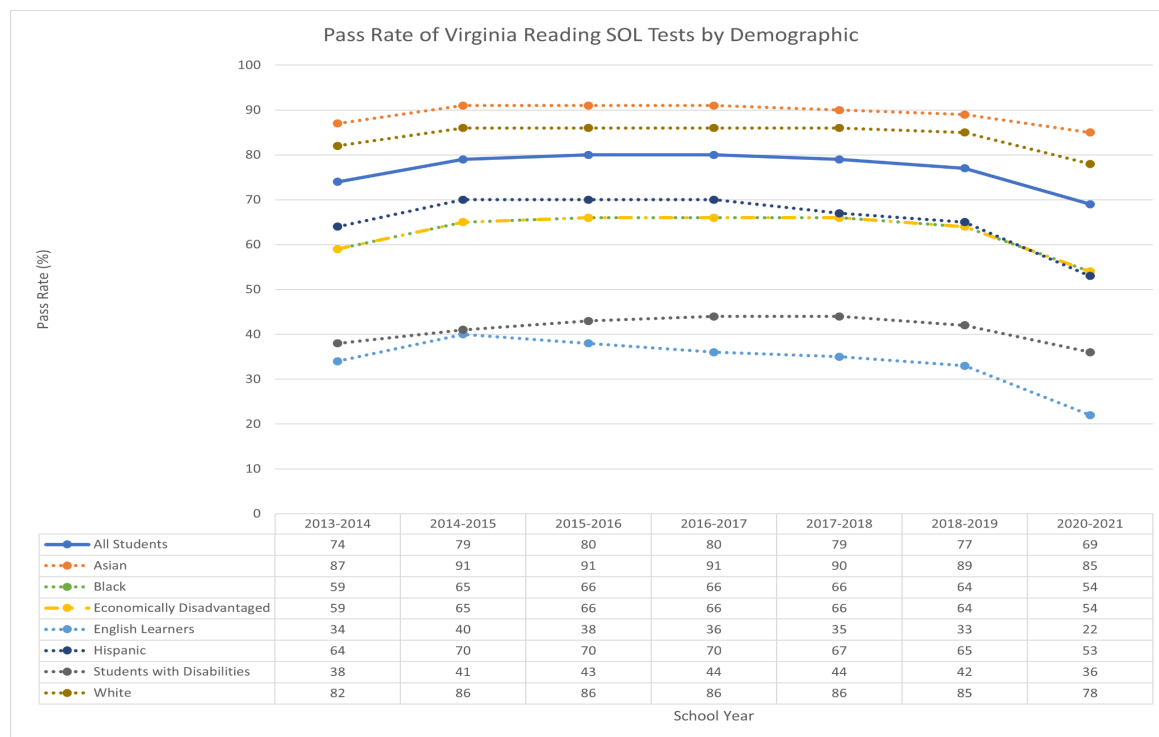|	2013-2014	2014-2015	2015-2016	2016-2017	2017-2018	2018-2019	2020-2021
All Students	74	79	80	80	79	77	69
Asian	87	91	91	91	90	89	85
Black	59	65	66	66	66	64	54
Economically Disadvantaged	59	65	66	66	66	64	54
English Learners	34	40	38	36	35	33	22
Hispanic	64	70	70	70	67	65	53
Students with Disabilities	38	41	43	44	44	42	36
White	82	86	86	86	86	85	78

As shown above, students' SOL scores have been in significant decline since the 2015-2016 school year. The most recent and significant of these drops was this past year 2020-2021. The pandemic has caused a gap in students' learning and disconnected the student from their school. Students would be more successful if they had an easy way to interact with and find their school's resources.

Teachers are struggling to get students invested in their own learning. Which is key to raising a child to be a lifelong learner. When students learn and what interests them and how they learn it empowers them in many aspects of their lives extending even beyond schooling.

Middle and high school students specifically are in the process of discovering themselves and providing a way for them to explore a possible interest in reading is important to our future. Students themselves are an investment in our future and it is important that they are given every opportunity to succeed and become happy and healthy learners.

Solution - AskMissy

Finding resources and books doesn't have to be as difficult as it currently is. Machine learning enables us to teach machines to do amazing things. Machine learning is already used by companies such as Netflix to find show recommendations so why not apply the same concepts to books? That is where AskMissy comes in. AskMissy seeks to use machine learning to produce personalized recommendations for students to get them more engaged in their own learning then they otherwise would be. Once students are more engaged in their learning it is likely that SOL scores will increase with that.

The age range AskMissy seeks to target is students in early middle school up to high school seniors. This age range is where SOLs are taken most often and where students would benefit the most from having an additional way to find books to read. Students

AskMissy will utilize machine learning to recommend books that students would engage well with based on their preferences, previous searches, teacher feedback, and their own feedback. The machine learning algorithm will take all of those factors into consideration when recommending a book to a student when they search.

Teachers and librarians are integral parts of the school system and students' learning. AskMissy does not seek to replace teachers and librarians as resources but to enhance their ability to help a wide range of students. Teachers and librarians will be an integral part of the design of the AskMissy program.

Product Description

Summary

As part of AskMissy's dedication to empowering both teachers and students AskMissy will implement a classroom feature that allows teachers to create class groups for students to join. This class group feature will allow teachers to give input on recommendations for students

as well as recommend books for the class as a whole. The teachers and students will be able to provide feedback on books recommended to provide more data for the machine learning algorithm to pull from.

AskMissy will be a web based application with user logins and user profiles that can be private or public based on user preferences. A user without a profile will only be able to access a basic search through the AskMissy and local libraries database. Once a user has created a profile they will have access to machine learning based search results. Which uses inputs from several sources to create results that will connect with the students.

There are two key machine learning algorithms involved in the AskMissy design. The first one is the language recognition algorithm. That algorithm is used to generate tags from book summaries gathered from the Goodreads API. These tags are then used as part of the second machine learning algorithm to connect users with tags that are in their preferences or they are likely to be interested in based on user feedback.

Goals

The goal of AskMissy is to improve students' ability to find supplemental educational materials. While striving for this goal we also seek to keep teachers and librarians informed on students' interests and educational journey so that they can better do their job in assisting students in their educational journey. As students become more engaged and invested in their own learning we expect SOLs to improve along with them. This is because reading SOLs are largely based on reading comprehension that would inevitably increase as students read more.

Key Product Features and Capabilities

AskMissy will be a web based application hosted online so users can have access to the program wherever they are. With the rise in smartphones and other smart devices it is important to develop a program that works on all types of internet devices. Users who do not have an

account will be considered guests and have limited access. However, users who do make

accounts with AskMissy will have access to the full capabilities of AskMissy. If student and

teacher users wish to connect their account with their school they will need to.

User Capabilities

Table 1:

Key: M = Manipulate A = Access							
Category	Feature	Guest	Student	Teacher	Librarian	Administrator	Tester
Data Retrieval	Metadata Report			M, A	M, A	M, A	
Data Retrieval	Basic Search	A	A	A	A	A	
Data Retrieval	AskMissy Search		A	A	A	A	
Data Management - Live Product	Machine Learning					M, A	
Data Management - Live Product	Source Tag Creation						
Data Management - Live Product	Source Tag Management				M	M	
Data Management - Live Product	Lesson Plans			M, A	M, A	M, A	
Data Management - Live Product	Internal Database Manipulation				M, A	M, A	
Data Management - Live Product	External Database Manipulation					A	
Data Management - Development	Source Tag Development						A
Data Management - Development	Machine Learning Training						A
Data Management - Development	Simulated Data						M, A
Data Management - Development	User testing reports						A

Category	Feature	Guest	Student	Teacher	Librarian	Administrator	Tester
Security	Login/Authentication	A	A	M, A	M, A	M, A	
	Data Encryption, moving					M, A	
	Data Encryption, resting					M, A	
Account Management	User Profile		A	M, A	M, A	M, A	
	Feedback		A	A	A	M, A	
	Group Management			A	M, A	M, A	
	Login/registration	A	A	M, A	M, A	M, A	
UI	Group Interaction		A	M, A	M, A	M, A	
	Bug Report	A	A	A	A	M, A	
	Basic Search	A	A	A	A	A	
	AskMissy Search		A	A	A	A	
	Communication		A	M, A	M, A	M, A	
	Personal Data Report	A	A	A	A	M, A	

Guests

Guest users will be using a limited version of AskMissy. They will not be able to access the main functionality of AskMissy being the personalized recommendation search. Without a proper account we would be unable to track preferences and collect input.

Students

Students who sign up for an account will be able to use the full functionality of AskMissy. They will get access to personalized recommendation search as well the ability to build a user profile that can be public or private. They also gain the ability to join classrooms and school systems provided they have a valid ID on file. They gain the ability to provide feedback to AskMissy and the school system. They do not have any manipulation privileges.

Teachers

Teachers who have verified accounts with a school system will be able to do all the same functions as students but also gain the ability to create and manage classrooms as well as receive feedback from the classroom.

Librarian

Librarians who have verified accounts will be able to receive and manage data from student feedback. Most importantly they will be able access and manage the database to update their local AskMissy library with new entrants.

Administrator

Administrators will be able to access and manipulate almost every aspect of the AskMissy program as to ensure proper use and smooth operation. Notable access to the database and machine learning algorithms will be important.

Tester

Testers are purely for development purposes only and are meant to simulate our most basic user, a student, while having access to the sample database to manipulate data to run appropriate tests with different sets of data.

Feedback

Users and teachers will be able to provide feedback to the AskMissy algorithm when they are given recommendations based on preference based searches. If the recommendation is or is not a good recommendation the feed back will good back into the algorithm and the algorithm will learn to avoid books with certain tags or recommend more books with certain tags.

Database

There will be several databases involved in the AskMissy system. Firstly there will be the school local database for their library which AskMissy will pull from to provide students with an inventory of books in the library. There will also be an AskMissy database which will contain user profiles, classrooms and outside resources. The final database will be the external API and database we will be using to get book information and summaries to put through the language recognition algorithm.

Machine Learning

There will be two main machine learning algorithms applied in the AskMissy design. Firstly the language recognition algorithm will be used to generate tags from book description pulled from an external API. These tags will then be attached to the book in the database and are used in the second machine learning algorithm. The second algorithm will be the resource matching algorithm. This algorithm will match users with recommendations when a user uses the AskMissy search functionality. The matching will be based on the tags generated in the first algorithm and user preferences.

Major Components (Hardware/Software)

The only user hardware that will be required to use AskMissy will be an internet capable device. As mentioned before, AskMissy will require three databases to work. The local library database, the AskMissy database, and the external API's database. The only database that will be kept and manipulated by the AskMissy team will be our database. We do not need access to manipulate any school or external database.

Since AskMissy will be built on and for the internet AskMissy will be utilizing HTML5, CSS3, and Javascript to create the frontend being the website itself. The backend algorithms will utilize python due to its strong machine learning and data management support. Such supports include Scipy and pytorch which will be integral to developing the prototype. The package manager we plan to use is anaconda to manage the code libraries the python code will run off of. The IDE we plan to use is Microsoft Visual Studio Code. For the repository and version management we will be using Gitlab due to the support it has with ODU servers. The database management tool we will be using will be mySQL.

Identification of Case Study

AskMissy is targeted at middle school and highschool students and their schools. AskMissy seeks to improve student SOL scores by enhancing student learning by getting students invested in their own education by providing them resources that connect with them. Getting a student interested in reading and learning would inherently increase SOL scores because the student would have a personal interest in learning rather than being forced to learn.

AskMissy will be used by school libraries to enhance their ability to recommend good books to their students. This is not meant as a replacement for teachers and librarians but as a supplementary material to help them get to know their students' interests better. The program does this allowing teachers to create classrooms and providing them with class statistics and classroom feedback.

As student interest in books grows, this product could also benefit local bookstores and book publishers. If AskMissy accomplishes its original intent it will grow the demand for books from the younger generation which is notoriously reading less.

Product Prototype Description

AskMissy innovates on typical book matching and recommendation services by using machine learning to enhance the user experience. The prototype will utilize user feedback to improve the machine learning as the prototype progresses. This is what allows the program to be innovative and unique. The AskMissy prototype will also implement safeguards to prevent users from accessing information not important to their role. So a student can have privacy if they do not wish to share their preferences.

The prototype will implement Guests, Students, Teachers, Librarians, Administrators, and Testers (Prototype only) within the program. Guests will have the most limited access and only be able to do basic searches. Students will be able to access the AskMissy search and fully utilize the program. Teachers will be able to manage classrooms and give input on recommendations for students. Librarians will also be able to give input on student recommendations while also managing the local database. Administrators will be able to address any issue within the database and the system in general. Finally, testers will be able to act as all of the roles and view the program from their perspective. This would enable them to effectively identify problems.

The search functionality will work in two different ways: the basic search and the AskMissy search. The basic search will only show results that relate to the title or author the user searched for. This search can also be used to search tags. The AskMissy search will enable users to unlock the full potential of AskMissy by producing machine learning based results. These results will be based on a combination of information from user feedback and teacher/librarian input.

The machine learning aspect of this prototype will utilize Association Rule Learning. This algorithm works to discover directional relationships between values. The algorithm will learn to find correlations between user preferences and books they enjoy. So in the AskMissy case the algorithm will identify when there is often common searches and common likes on a book and strengthen the recommendation power of that book for those search preferences.

There will be no special hardware utilized for this prototype. The only hardware that will be utilized to develop the prototype will be the virtual machine provided by the school and any internet connected device. For the software we will be utilizing Git, Flask, PyCharm, and Docker. Git will be used for version control and task management. Flask will be used to develop the user interface and develop the user experience as a whole. PyCharm will be used to write the python code that will be used for machine learning. Docker will be used in the VM to package the software within the VM and help us create the web server on the VM.

Glossary

Administrator: A user who is responsible for managing a majority of AskMissy's working data.

Agile: A set of frameworks and practices where solutions evolve through collaboration between self-organizing cross-functional teams.

AskMissy: A software application that will help users find more relevant resources.

Application Programming Interface (API): A software intermediary that allows two applications to talk to each other.

Data Retention: The continued storage of an organization's data for compliance or business reasons.

Database: An organized collection of structured information, data, typically stored in a computer system.

Economically Disadvantaged: A student eligible for Free/Reduced Meals who receives Temporary Assistance for Needy Families (TANF) or is eligible for Medicaid.

Exact Match Search: A search for a single specific type of resource.

File Server: A device that controls access to separately stored files.

Guest: A user who is not a student, teacher, librarian, or administrator; who has limited access to the AskMissy program.

Librarian: A user responsible for managing the library's inventory/database, communicating with teachers and students.

Personal Learning: An educational approach that aims to customize learning for each user's strengths, needs, skills, and interests.

Student: A user studying at a K-12 education institution in need of reliable resources.

Teacher: A user who helps K-12 students acquire knowledge. They are responsible for making plans and managing students' groups/communication.

Temporary Assistance for Needy Families (TANF): A program that provides eligible families with a monthly cash payment to meet their basic needs.

Tester: A user responsible for designing and conducting testing suites for usability

testing. **User:** An individual using the AskMissy Interface.

Web Scraping: The process of extracting content and data from a website. **Web**

Server: A computer program that distributes web pages as they are requisitioned.

Windows: A series of operating systems developed by Microsoft.

References

Albanese, A. (2021, May 6). *Report urges library leaders to address decline in public library usage stats*. PublishersWeekly. Retrieved December 15, 2021, from https://www.publishersweekly.com/pw/by-topic/industry-news/libraries/article/86256-report-urges-library-leaders-to-address-decline-in-public-library-usage.html

"Allreaders.com features detailed book and movie reviews from many different genres of books!," *Detailed Book review summaries*. Retrieved September 22, 2021, from http://allreaders.com/.

"Amazon.com: Kindle eBooks: Kindle Store: Nonfiction, Literature & Fiction, Foreign Languages, Business & Money & More," *Amazon*, Retrieved October 01, 2021, from https://www.amazon.com/Kindle-eBooks/

Coates, T. (2021). *Freckle report 2021: Digital or diverse?- the future for public libraries*. Tim Coates Books.

Crain, C., & Waldman, K. (2018, June 14). *Why we don't read, revisited*. The New Yorker. Retrieved September 22, 2021, from https://www.newyorker.com/culture/cultural-comment/why-we-dont-read-revisited.

Ellard, C. (2020). *Covid-19 impact on Sol Assessments, verified credits, and graduation requirements for Spring/Summer 2020*. Newport News Public Schools. Retrieved December 14, 2021, from http://sbo.nn.k12.va.us/sol/covidimpact.html

LAB 1 – ASKMISSY PRODUCT DESCRIPTION 14

Fuglei, M. (2019, July 22). *Why students who read for pleasure are stronger academically.* ResilientEducator. Retrieved October 1, 2021, from

https://resilienteducator.com/classroom-resources/how-reading-for-pleasure-helps-studen

t s-develop-academically/.

Garcia, E., Weiss, E., & Welshans, I. (2020, October 7). *What teaching is like during the

pandemic-and a reminder that listening to teachers is critical to solving the challenges

the coronavirus has brought to public education*. Economic Policy Institute. Retrieved

October 7, 2021, from

https://www.epi.org/blog/what-teaching-is-like-during-the-pandemic-and-a-reminder-that

-listening-to-teachers-is-critical-to-solving-the-challenges-the-coronavirus-has-brought-t

o -public-education/.

GeeksforGeeks. (2020, August 29). *5 most-recommended career fields in computer

science*. GeeksforGeeks. Retrieved January 25, 2022, from

https://www.geeksforgeeks.org/5-most-recommended-career-fields-in-computer-science/.

Gelles-Watnick, R., & Perrin, A. (2021, September 21). *Who doesn't read books in

America?* Pew Research Center. Retrieved September 22, 2021, from

https://www.pewresearch.org/fact-tank/2021/09/21/who-doesnt-read-books-in-america/

. Gioia, D. (n.d.). Reading at Risk. Washington D.C., Virginia ; National Endowment for the

Arts. https://www.arts.gov/sites/default/files/RaRExec_0.pdf

Henry. (2021, May 27). *The importance of Reading* Retrieved September 22, 2021, from

https://www.uopeople.edu/blog/why-its-important-to-read/

Ingraham, C. (2018, June 29). *Leisure reading in the U.S. is at an all-time low*.

The Washington Post. Retrieved September 22, 2021, from

https://www.washingtonpost.com/news/wonk/wp/2018/06/29/leisure-reading-in-the-u-s-

i s-at-an-all-time-low/

"Meet your next favorite book," *Goodreads*. Retrieved October 01, 2021, from

https://www.goodreads.com/

Frequently asked questions about Sol testing. Virginia Department of Education.

(2015, October). Retrieved October 7, 2021, from

https://www.doe.virginia.gov/testing/sol_faq.pdf.

Rosalina, E. N. (2018, November 30). *The correlation between self-esteem and student's*

reading comprehension. English Language Teaching Educational Journal. Retrieved

October 1, 2021, from

https://eric.ed.gov/?id=EJ1283078

What should I read next? book recommendations from readers like you. What Should I Read

Next? (n.d.). Retrieved January 25, 2022, from

https://www.whatshouldireadnext.com/

Virginia Department of Education. (n.d.). *Sol Test Pass Rates & other results*.

Retrieved October 7, 2021, from

https://www.doe.virginia.gov/statistics_reports/sol-pass-rates/index.shtml.