Chapter 1

Introduction

With the development of technology more and more focus is shifting on teaching kids the concepts of programming. Online learning is becoming an important part of our modern world. Development of Creative thinking is one of the fundamentals of early education. While kids usually get excited about learning new things, they lose interest pretty quickly. As difficult and challenging as the subject of programming could be, keeping the audience interested is even more so. Scratch by LLK group at MIT media labs is one such platform that aims developing programming skills in kids while also keeping them interested. Scratch is not only aimed at developing programming skills but also enabling Creative thinking and collaborating skills in children [1].

Scratch is a visual programming language/ environment that enables the user to create rich interactive projects including various media [2]. We can see various examples where users have created a wide range of projects like animations, storylines, games, music videos, online games, science projects, and sensor-driven art (Figure 1).

The Scratch platform lets user import media like sound and images or creates using inbuilt editor, these objects ( called Sprits ) can be made to interact via programming. Programming is done using lego-like blocks that snap together to form a logical command. This projects can then be saved on the computer or can be shared online on the scratch website.

A key design goal of Scratch is to support self-directed learning through tinkering and collaboration with peers [2]. Many users learn Scratch as they go, trying commands from the palette or exploring code from existing projects. To encourage such self-directed learning, the Scratch programming environment was designed to invite scripting, provide immediate feedback for script execution, and make execution and data visible [2]. When a project is shared online on the scratch website, it can be seen, other peers/users. Then the project can then get comments by peers. Peers can also use the existing project and build upon the top of it. This is known as remixing. Scratches encourage users to explore the work of others and collaborate to build on it. At the beginning some scratchers complained about the projects remix, claiming that others were stealing their work [45]. This led to a discussion on the website forum about the idea of sharing in an open source community. Scratches goal is to make the users feel proud when their projects are remixed and to get upset. Scratch continuously adds new features to the website to encourage this attitude. When a project is remixed there is a link added to the original project giving the original author the due credit. Also top remixed projects are featured on the websites homepage.

Some interesting use of scratch is the users formign online “companies”.Some scratchers founded online “companies”. A 15 year old girl from England who goes by the screen name BeeBop, set up a project with animated sprites and encourage users to share them in their own project or give a special order on request for a custom made sprite. She was starting up the no fee consulting company. Later a 10 year old girl also from UK with screen name MusicalMoon, liked the project by BeeBop and asked if she would be interested in making a custom background for her own project. This is how Mesh Inc. started a self declared “mini company” which provided remarkable games in Scratch. After some time the company was joined by a 14 year old from New Jersey by the screen name Hobbit, who offered his service styling “I am a fairly good programmer and I could help with debugging and stuff .[45].

Scratch is also being used in formal learning environments at universities to introduce children and high-school students to programming concepts. Studies have shown that visual programming concepts have a notable impact on the learning of computer programming concepts [26]. Some of the well-known visual programming languages like Alice [27], RAPTOR [28] and Etoys[29] have been used by different educational institutes. Scratch possibly is the most popular and widespread visual programming language today [30] [31]. Although originally designed for kids, it was accepted for introductory programming lessons by a few higher education organizations [32].

1.1 Motivation

The development of Scratch is always tightly associated with the development of Scratch website [4]. For Scratch to be successful, it must be connected to a community which supports collaboration and feedback from one another and improving on others work [44]. Thus the concept of sharing is inherently built into Scratch with a “Share” option at the top of the screen.Clicking the button uploads your project to the community website. Once the project is shared on the website anyone can run it , tinker with it comment on it and vote using the Love button.

Following the launch of Scratch in 27 months, users have shared more than 500,000 projects on the community website [45].For many Scratch users the concept of sharing their work with others and getting feedback and advice from peers is the source of strong motivation. The large amount of work shared on Scratch serves as a library for exploring new ideas and learning new techniques in programming.

Scratch website also serves as a ground for collaboration. Many Scratchers are constantly adapting and building on one another's ideas. Around 15% of the total projects shared on scratch are remixed projects meaning they were built upon some other project that some other Scratcher has previously shared [45]. For instance there are a number of different versions of Tetris, where different Scratchers have added new features to the game to try and improve the experience. There are also many dress up-doll projects , contests which are all modeled after previous Scratch projects.

One of the main objectives of Scratch is to encourage collaboration amongst peers [2]. This is done by sharing the projects online with the community, getting feedback, tinkering with and mixing others work with their own ideas. If a user likes the project scratch provides options to “LIKE”, “LOVE” or “REMIX” it, Users can also leave comments on the project. Users can add other users ( for example the author of the project) as friends i.e. follow them.

The ways of discovering other users work which has been shared is done on the scratch website. On the main home page are the sections like *Featured Projects, Featured Studios, Projects Curated by, Scratch Design Studio - Puzzled, What the Community is Remixing, What the Community is Loving,* This sections features projects that are either hand-picked by curators, or what are most trending projects in the community or the projects that are getting most attention(likes, loves, remixes etc).

However, when it comes to personalised recommendation there is no option which looks at what the user is interested in. Like in any online platform self-directed learning plays an important part in scratch. For enabling suitable self-directed learning it is important to guide the user through proper steps and resources [3]. By recommending projects based on users development we can better enable user with self-directed learning.

Thus this study focuses on developing an effective recommendation system for Scratch which would recommend users with the personalised recommendation, based on their preference and development.

1.2 Research Objective

With the research goal in mind for “Studying recommender system in Scratch platform,” this dissertation focuses on the following objectives

We will study and apply various collaborative filtering and content-based filtering models on the Scratch dataset. The dataset made publicly available by the LLK Scratch team will be used for analysis and model building. We will try to fit an item-item collaborative model on the project datasets and compare the recommendation outputs using standard metrics. We will also discuss heuristics-based recommendation and hybrid recommendation method.

We will also demo how a recommendation on scratch would work on a scratch website by building a JS client/server model. Finally, we discuss the risks and limitations of recommendations methods.

1.3 Research Challenges

1. To filter the huge dataset available for finding the relevant and useful information.

2. Find the best collaborative filtering algorithm that fits the available dataset.

3. Finding the best metrics for Evaluating the generated recommendations for quality and usefulness.

1.4 Dissertation Overview

Recommender system are all around from the product recommendation you get in Amazon to Spotify generating a customized playlist.This dissertation discusses the Collaborative filtering technique which is one of the most common ways in recommender system area for calculating the predictions of items for a given user. There are many different techniques in collaborative filtering which can be applied depending on the requirements and the data. For predicting the Scratch projects a user might find interesting in Scratch, item - item collaborative filtering techniques is best suited. In item-item collaborative filtering, we first calculate the closeness of the project to other projects from our dataset, and then check what projects from this list the user is most likely to like.This is done by looking at the past activities of the user.

The data is provided by the Scratch community website. It contains details of users, projects and the project contents that have been shared on the website. The data is collected over a period of 5 years from 2007 to 2012. The data is then filtered to see for the attributes that will be useful in our task for building a recommender system. We find that the most useful from the dataset is the users and the projects table, which contains enough information for our recommendation system analysis.

After fitting the item-item collaborative filtering model, for the purpose of demonstrating the recommendations system, a client/server model is built. JavaScript and Python have been used for the client-server interaction and Pandas library for data analysis and calculating the similarity matrix.

1.5 Dissertation structure

The rest of this document is orchestrated as following. Chapter 2 discusses the current state of the art in the field of recommender systems and touches the topic of recommendation system in education. In chapter 3, we apply the information filtering techniques discussed previously on the public dataset for Scratch and study the outputs. Chapter 4 discusses the evaluation matrix and limitations of this dataset and the methods. Finally, in chapter 5 we conclude with the scope of future work.