yourtutor

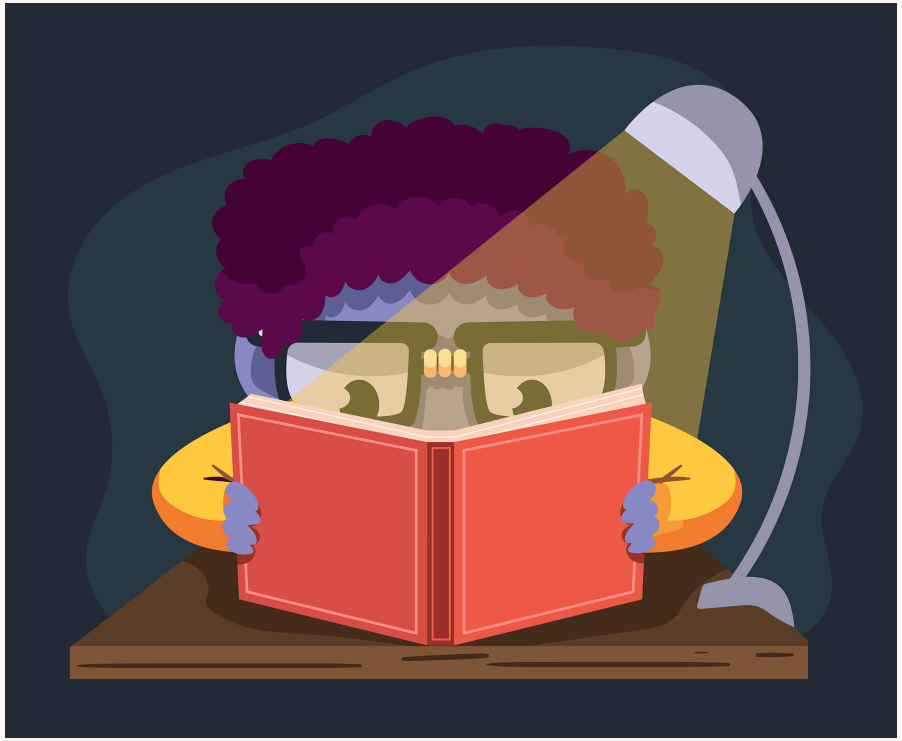
by **BrAIn Power**

HackYeah 2020

Problem Statement

The COVID-19 pandemic resulted in close-down of schools in whole countries, leaving students without teachers and psychological support right before critical exams on their education path. Media is flooding with problems that students face with trying to attend classes via videoconferences, lack of appropriate preparation for maturity exams and uncertainty about their nearest future, causing **stress, anxiety** and **delay in education.**

Videoconferences for classes do, in some extent, solve the problem as they allow for conducting classes. However, moving a concept of a group of students listening to a teacher online does not utilize the potential of technology in online learning for schools.



Problem Solution

We present a general framework which makes use of technology to **improve teachers' efficiency and optimize students' time** by introducing a personalized self-learning system.

Our simulation shows abilities of **Reinforcement Learning** algorithm to be useful in teaching. A virtual teacher assesses student's level in a defined area (such as mathematics, literature) by having them solve tasks of defined difficulty and automatically adjusts the difficulty level of following tasks by measuring the time of learning which student needed in order to obtain certain skills.

Moreover, by using appropriate measuring functions, the virtual teacher can suggest moving to other areas so that the student does not focus on a single subject.

Benefits

- The presented framework is **general** and can be applied to **any environment** in which someone is improving over time by solving tasks and performing tests of their skill.

- It optimizes students' time as it allows for students with predispositions in certain areas to **proceed faster with more difficult tasks**, and for students which are struggling to **spend more time in areas they find challenging**.

- It improves teachers' efficiency as they can observe which tasks are **the most challenging for students.** It also helps in **identification of talents** and students who need **extra help**.

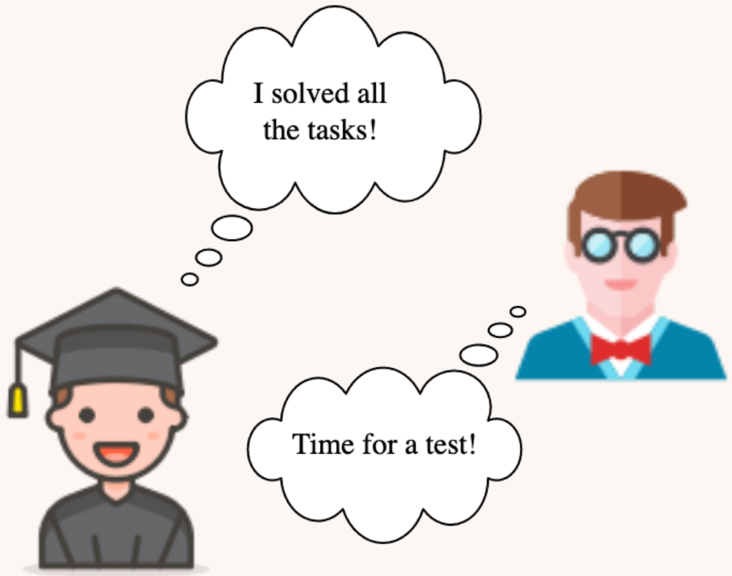
- In times of crisis it may bring **psychological comfort** for both students and teachers. Students are guided towards improving their skills before important exams and teachers can perform with system's support.

Environment Model

In our simulation we use **Reinforcement Learning** in which an **agent** representing a **teacher**, learns itself to make decisions which result in optimized learning of **each student**.

Students are treated as environment with which the **agent** interacts. They are modeled with randomly sampled expertise and efficiency levels in defined areas, which define their **hidden states**. It means that one student can be on medium level and improve slowly, while other may start from easy level and improve rapidly. **The hidden states are unknown to the teacher.**

Example of modeling interpretation:
Student A has **high level** of knowledge in **mathematics** and their test result **improves by 1%** on average after **30 minutes** of learning.



Simulation Flow

A step is defined as a set of moments when student is awaiting teacher’s decision.

At the beginning of each element in step, teacher can make one of the following decisions:

- advise student to **learn**. Learning can be performed for any **subject, difficulty level** and for **amount of time**. Set of possible decision contains all possible combinations of those variables.
- advise student to **solve a test**. A test can be solved for any **subject, difficulty level** and for **amount of time**. Set of possible decision contains all possible combinations of those variables.

Having received the decision from the teacher, the student **increases** their expertise in subject depending on its difficulty and own hidden state after learning. After **testing,** the student responses with % test result based on their expertise and test’s difficulty level.

Reward Mechanism

The reward system for teacher followes rules:

- Penalty for every testing and learning, the longer process the bigger the penalty
- Receives reward or penalty based on comparison of test result with previously received test result
- Receives a big reward for score in test over 90%

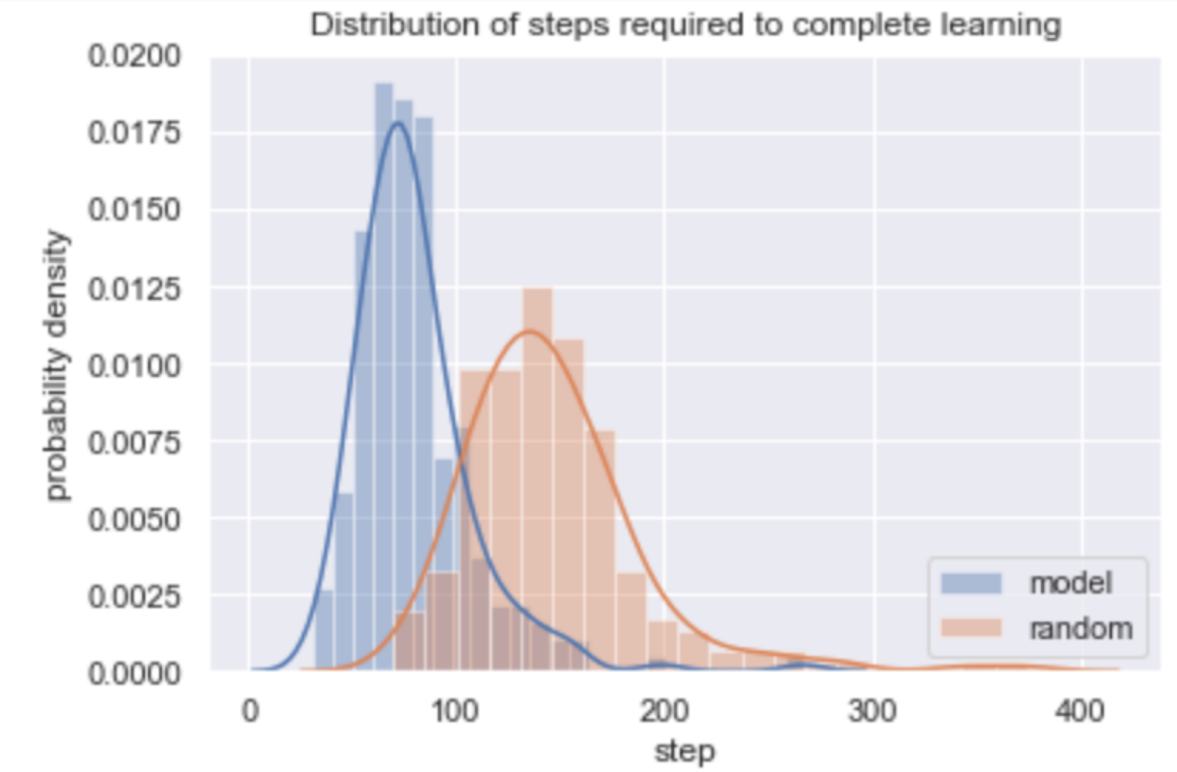
Such model minimizes learning time and maximizes expertise increase.

Observations

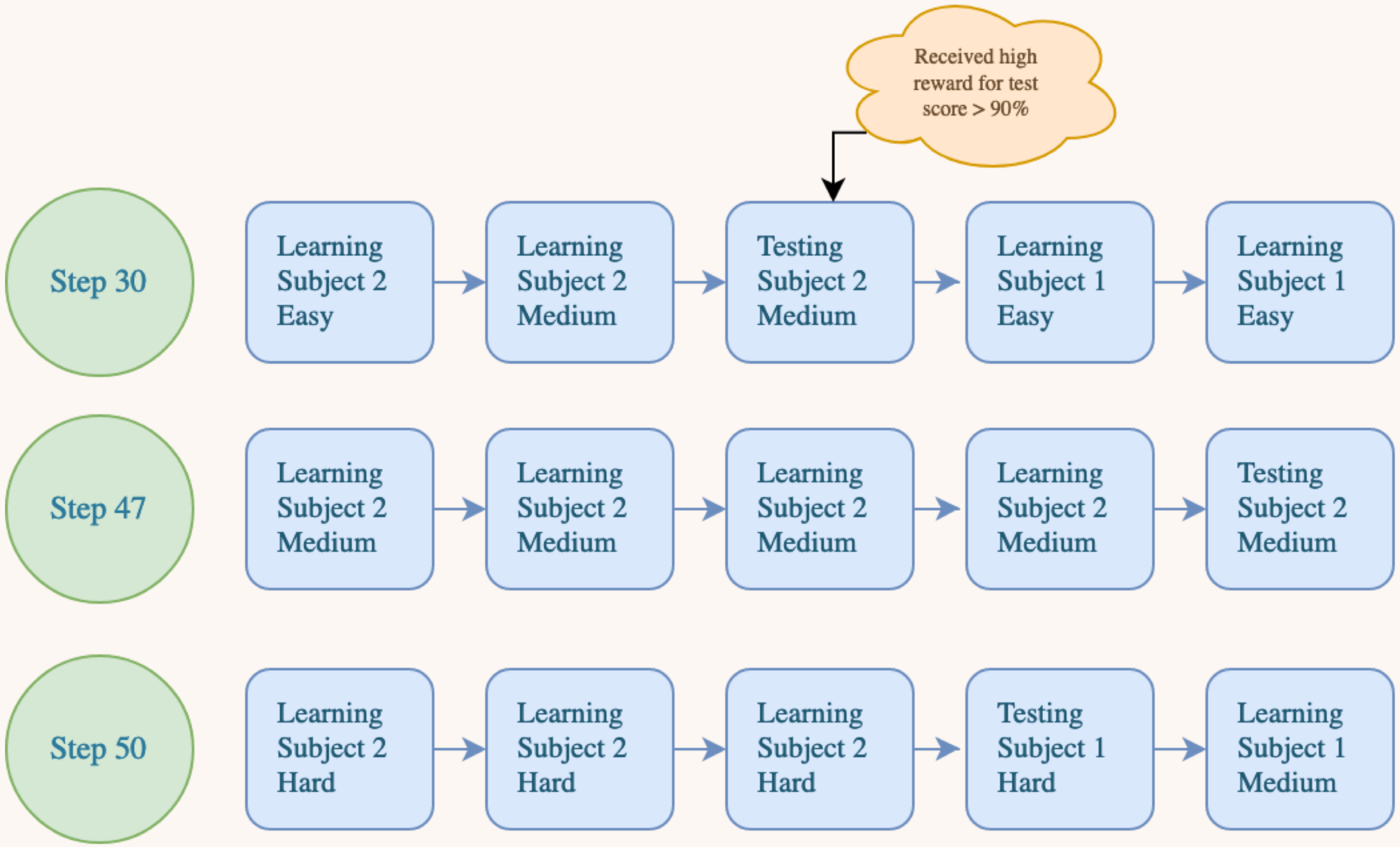
During the simulation we observed that as the training progressed, the teacher was making **more and more reasonable decisions** in regards to its reward strategy.

On the right we can see a distribution of number of steps needed to complete learning (achieve over 90% on a test) for our simple model and random decision. **The fact that such simple environment works well is very promising!**

We also performed experiments with reward strategy in which the teacher gets no reward for learning – after some time it realised it has to **pay the price of learning to achieve success** in testing!



Sample decisions



In the presented examples we can see the agent performing reasonable choices.
In **Step 30**, the teacher directs student to another subject after they received high score in test for one subject.

In **Step 47** the agent probably learned that this student needs more learning time in this subject.

In **Step 50** we can see that the estimated skill was higher than student's actual expertise, thus the level of learning is decreased.

Applications and enhancements

- The framework can be applied to any learning environment: be it online learning for school or gaining professional experience.

- The framework could be useful in recruitment processes for **assessment of candidate's learning abilities**.

- It can be easily supplemented with intelligent task checking, such as automatic essay evaluation.

Success of the teacher is determined by the quality of the students model. The student is a crucial component of the framework. A good model representing a student could be achieved by **cooperation with educators and psychologists** to address key factors in education process.

Knowing possibilities of each student, system using this framework could anonymously pair students with high and low expertise to build **feeling of community** and **responsibility**, as well as for learning optimization.

Technology + Repo

Simulation was performed using **Open AI Gym** environment and usage of Reinforcement Learning:

- ❖ Proximal Policy Optimization 2
- ❖ Trust Region Policy Optimization

Code for the project can be found here: https://github.com/Ewande/hackyeah2020/tree/environment

Thank you!