AI Helps Duolingo Personalize Language Learning

Learning a foreign language was probably one of your goals last year. And the year before, and the year before that. Like gym memberships, our best intentions often don't survive very long. Aside from the time required to achieve proficiency with a new language, most people struggle with traditional approaches to learning. Even many web-based language tools can be monotonous and cumbersome.

Pittsburgh-based startup Duolingo is changing all that with its AI-based language-learning platform. The company reaches over 300 million users with more than 32 language courses—from French and Tamil to endangered languages such as Hawaiian and Navajo.

What's different about Duolingo is its personalized learning approach, which gamifies the learning experience through a points-based reward system to keep users engaged and progressing toward proficiency. The U.S. State Department estimates it takes 600 hours to learn a category-one language like French or Italian. Duolingo hopes you can do this in just 15 minutes a day.

Users start with Duolingo's AI-driven adaptive placement test which probes them with real exercises they would take during the course—so if you've had four years of high-school French, you don't have to start at the beginning of the most basic course. Each question or challenge in the test is adaptively chosen based on the previous question, and whether you got it right or wrong.

"The difficulty of the words, the grammar, and the way we present it to you in the test, all play a role to pick the exact configuration so that in less than five minutes we have a really good sense of where you're going to start the course," explains Burr Settles, Research Director at Duolingo.

Using a concept called spaced repetition, language lessons are designed so users practice personalized tasks over longer and longer intervals, which has proven to be more effective than cramming in a short period of time.

As you become more proficient, you interact with the content in different ways. For example, for each word in the curriculum, Duolingo keeps track of how many times you've seen it, how many times you've gotten it correct, the modes under which you got it correct, and how long it's been since you've practiced it.

to keep practicing, exactly when you need it."

"We can inject what you need to keep practicing, exactly when you need it."

Burr Settles
Research Director
Duolingo

The learning behind the lingo

To enable this AI, Duolingo uses deep learning, a subset of AI and machine learning that uses neural networks to mimic the brain's behavior to quickly analyze data and make intelligent predictions. Using deep learning algorithms for natural language processing, the company can analyze user log data to predict the likelihood that users will get an answer correct. These predictions are the basis for personalizing both the adaptive learning test and content for the learning app.

users a foreign language by having them translate documents such as articles from Wikipedia or news sites. At the time, Monolingo (and even early Duolingo) used more traditional cognitive science algorithms. For example, baseline algorithms used handpicked parameters—meaning they weren't necessarily learning from real data. As Duolingo researchers A/B tested various approaches with users, it was clear that for the level of personalization they were targeting, more sophisticated and customized machine learning models were necessary.

"These are very tailored problems, so we've had to invent everything from the ground up," says Burr. "It's a normal lifecycle for these use cases to try a rudimentary cognitive approach first to start collecting data, and then once you have data, start refining it with deep learning."

To develop these custom algorithms (everything from non-native speech recognition to classification for automated scoring), Duolingo uses the PyTorch deep learning framework on Amazon Web Services (AWS). These deep learning models are trained and then deployed into production using Amazon EC2 P3 high-performance GPU instances. Speed and scalability are essential for training since the models may use anywhere from 100,000 to 30 million data points at a time, depending on the problem, in order to make more than 300 million predictions every day.

"We'll use a sliding window because just two weeks of data is plenty given the number of users, number of tests, number of languages, to train our models," says Burr. For managing data pipelines for machine learning, the company uses Amazon DynamoDB for data management, Amazon EMR with Amazon EBS as temporary storage, Amazon S3 for permanent storage, and Spark to perform computations for periodic batch predictions.

As well, to bring its applications to life, Duolingo uses Amazon Polly, a deep learning-powered text-to-speech tool that easily integrates into its applications, to give voice to the test and to numerous courses.

By using these deep learning tools, the company saw an improvement in both prediction accuracy and user engagement. The number of users who used Duolingo and came back the second day improved immediately by 12 percent.

context. For example, you may get a question wrong but why you got it wrong is not as clear. Did you get it wrong because you forgot the word? Or maybe you conjugated it wrong.

"It's not always clear to tease out from the signal we get back what the cause was," says Burr. "There's a lot more AI to do."

As Duolingo improves its language offerings using deep learning, it's possible you'll be able to strike at least one goal off your New Year's resolutions.

Related Stories

Kia uses ML to reduce fatalities from auto accidents

Learn more »

TuSimple uses ML to train self-driving semis

Learn more »

🔍 ☰

GE Healthcare drives better healthcare outcomes with ML

Learn more »

Zocdoc uses AI to build patient confidence

Learn more »

Sign In to the Console

Learn About AWS

What Is AWS?

What Is Cloud Computing?

What Is DevOps?

What Is a Container?

What Is a Data Lake?

AWS Cloud Security

What's New

Blogs

Press Releases

Resources for AWS

Getting Started

Training and Certification

AWS Solutions Portfolio

Architecture Center

Product and Technical FAQs

Analyst Reports

AWS Partner Network

Developers on AWS

Developer Center

SDKs & Tools

.NET on AWS

Python on AWS

Java on AWS

PHP on AWS

Javascript on AWS

Help

Contact Us

AWS Careers

File a Support Ticket

Legal

<div style="border: 1px solid orange; padding: 10px; text-align: center;">

Complete Sign Up

</div>

Amazon is an Equal Opportunity Employer: *Minority / Women / Disability / Veteran / Gender Identity / Sexual Orientation / Age.*

Language

عربي |

Bahasa Indonesia |

Deutsch |

English |

Español |

Français |

Italiano |

Português |

Tiếng Việt |

Türkçe |

Русский |

ไทย |

日本語 |

한국어 |

中文 (简体) |

中文 (繁體)

Privacy

|

Site Terms

|

Cookie Preferences

|

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.