MODULE 1 UNIT 1
Lesson Video 1 Transcript

IN COLLABORATION WITH getsmarter™

Module 1 Unit 1 Lesson Video 1 Transcript

Practical application of machine learning

KOSTA KALOGEROPOULOS: Machine learning is all around us, from Apple's FaceID to unlock our phones, tracking our fitness with a Fitbit, Google Photos arranging our photos into categories, and even Facebook's algorithms detecting people from the photos we upload. Some of these applications of machine learning may be obvious, but there are also a number of applications that may not be that obvious, like the spam folder in your email account. In this unit, you were introduced to the concept of machine learning, taking a closer look at the definition and the different types of machine learning that exist. Now that you understand what machine learning is, let's look at a practical example of machine learning you probably encountered recently.

The problem statement: Email spam folders

KALOGEROPOULOS: We are all too familiar with the unwanted advertisements, chain letters, and other promotional material sent to us by organisations getting hold of our personal information. Instead of filtering through our emails manually, machine learning has helped us by identifying these emails as they are sent to us, and arranging them all into a spam folder automatically. If it wasn't for the occasional incorrect classification by the machine learning algorithm that classifies an online receipt incorrectly, we would never ever interact with that folder at all.

Practical application of email spam detection

KALOGEROPOULOS: Since the machine learning model aims to classify an email into two categories, 1) spam or 2) not spam, this is a classification problem. The way that the machine learning model does this is by scanning the email and finding the words or characters that frequently appear in spam emails, indicating that it might be spam. Most of the attributes indicate whether a particular word or character has frequently appeared in the email, for example, the length of sequences of consecutive capital letters.

The machine learning would classify incoming emails in the following way. Firstly, by inspecting the words, characters, or their combinations in the email. Secondly, access a library of existing emails – both genuine and spam, i.e. the training data – then assessing how likely it is for these words to have originated from spam emails, and finally suggesting a classification: spam or not. Of course, the machine learning sometimes gets it wrong, sending non-spam emails to the spam folder and vice versa. This can happen either due to model assumptions on how to use or combine words and/or the quality of the training data, e.g. whether the training data contain enough patterns of text in spam email.

Conclusion

KALOGEROPOULOS: In this video, you explored a practical example of a classification problem in machine learning, and how these models are used to classify emails into spam folders. This is only one of the many examples where machine learning is applied to solve problems, not only for individuals, but also for businesses. Now that you have a better understanding of how machine learning is practically applied, engage with the rest of the lesson to learn more about other practical applications of machine learning and how these

applications contribute to business productivity, possibly generating revenue. What would you like to see next in the space of email classification? Are there any other machine learning models you would like to develop for your organisation to classify emails into specific folders?