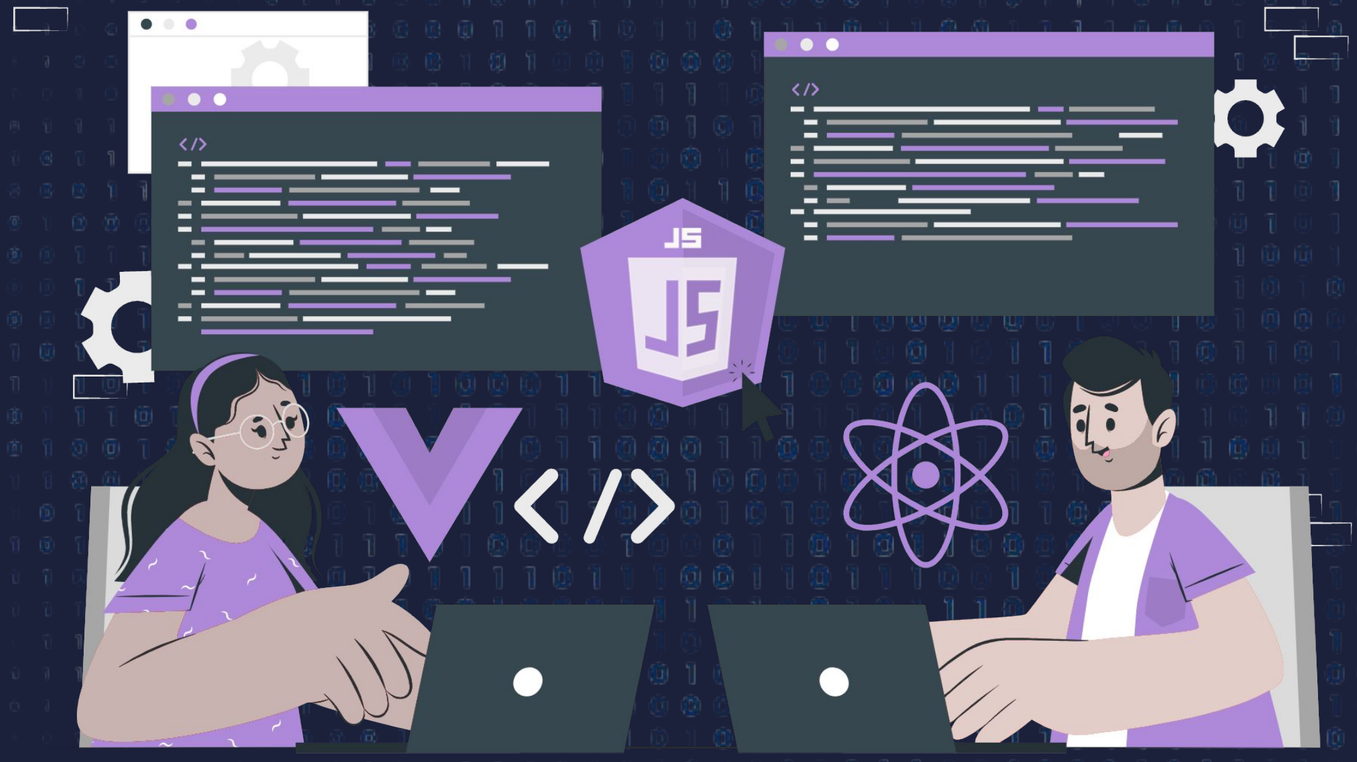
PW SKILLS

Writing regex with example



Lecture CheckList

1. Introduction.
2. Simple examples of regex.
3. Regex for email validation.

Introduction

From the previous lectures, we have a clear idea of what regex is, how to use regex, and the applications of regex. In this lecture let's look at some examples where we will be writing regex.

Let's start with some simple examples of finding a string pattern. To write regex we first need a pattern which we need to search for and one more important part of regex syntax is the "/pattern/". The pattern inside the slashes is the search pattern you want to match in a string.

Simple examples of regex.

Regex for email validation.

Till now we have seen the basic implementation of regex. Now, let's look at some real-life challenges like using regex to validate email addresses.

Using regex to verify an email address. An email is a string separated into two parts by @ symbol. The first part contains personal information while the other contains the domain name at which the email is registered.

Regex for email validation.

The personal information can contain:

- Uppercase and lowercase letters (A-Z and a-z)
- Numeric characters (0-9)
- Special characters - ! # $ % & ' * + - / = ? ^ _ ` { | } ~
- Period, dot, or full stop (.) with the condition that it cannot be the first or last letter of the email and cannot repeat one after another.

The domain name contains:

- Letters
- Digits
- Hyphens
- Dots

Regex for email validation.

First, let's write the pattern to be searched for. We can break down the regex pattern as follows:

- ^: This is an anchor that matches the start of the string.
- [a-zA-Z0-9.!#$%&'*+/=?^_{|}~-]+`: This matches one or more characters from the specified character set. The character set includes letters (upper- and lower-case), digits, and special characters that are often used in email addresses.
- @: This matches the at symbol.
- [a-zA-Z0-9-]+: This matches one or more characters from the specified character set, which includes letters (upper- and lower-case) and digits.
- (?:\.[a-zA-Z0-9-]+)*: This matches zero or more instances of a sequence of a dot followed by one or more characters from the specified character set.
- $: This is an anchor that matches the end of the string.

THANK YOU