

Interpretabilidade de Modelos de Machine Learning focados na predição de defeitos em software

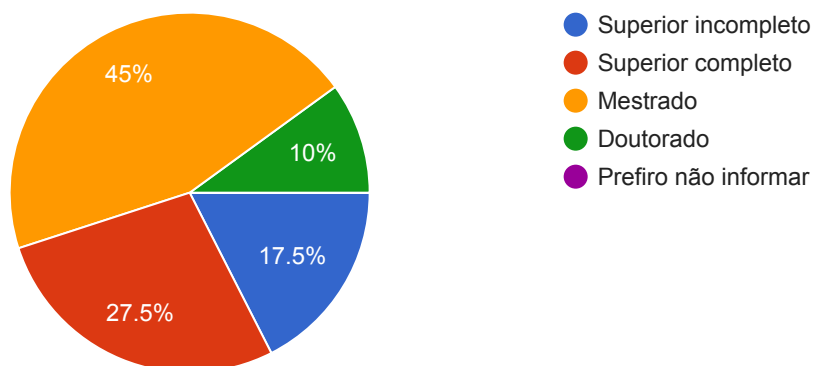
40 responses

[Publish analytics](#)

Perfil dos candidatos

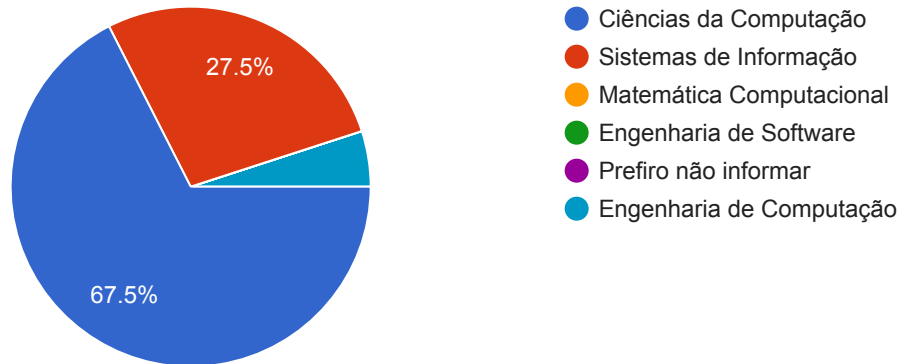
Qual sua formação acadêmica na computação?

40 responses



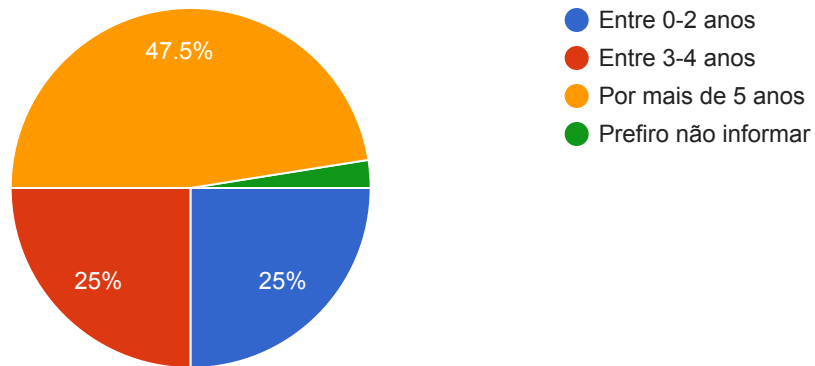
Em qual curso você é graduado (se tiver superior incompleto, indique o curso que começou) ?

40 responses



Há quanto tempo você desenvolve software?

40 responses



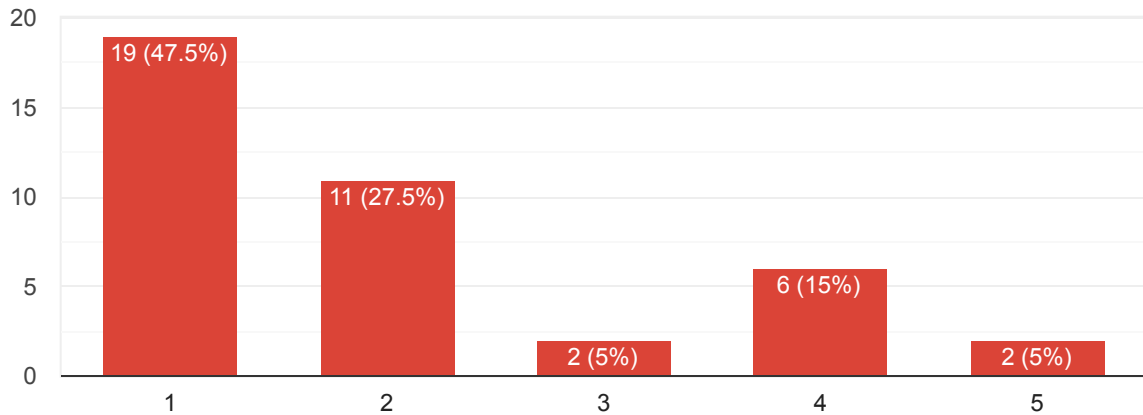
Estudo sobre a Interpretabilidade

Primeiramente, apresentamos 5 features que serão usadas na explicação local sobre o modelo:



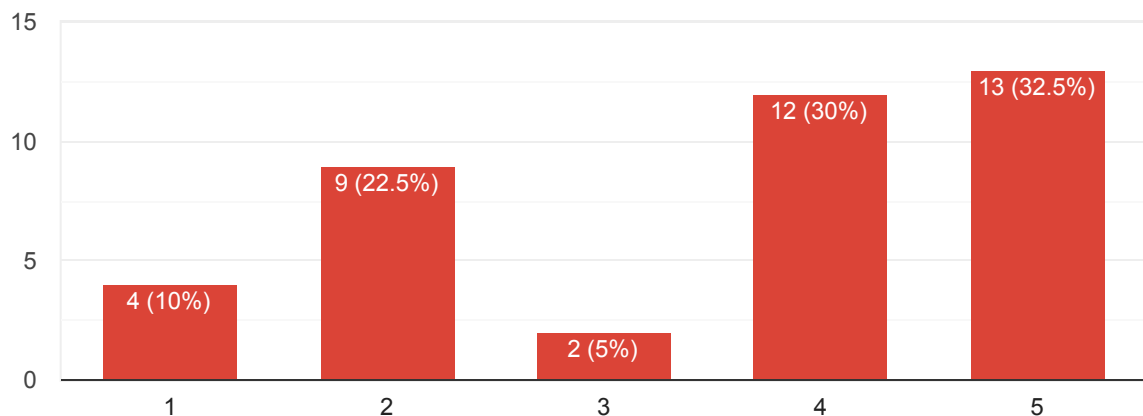
Considerando a explicação local do módulo de software apresentado acima, você consideraria aumentar o LOC_TOTAL, ou seja, aumentar as linhas de código do módulo para evitar defeitos?

40 responses



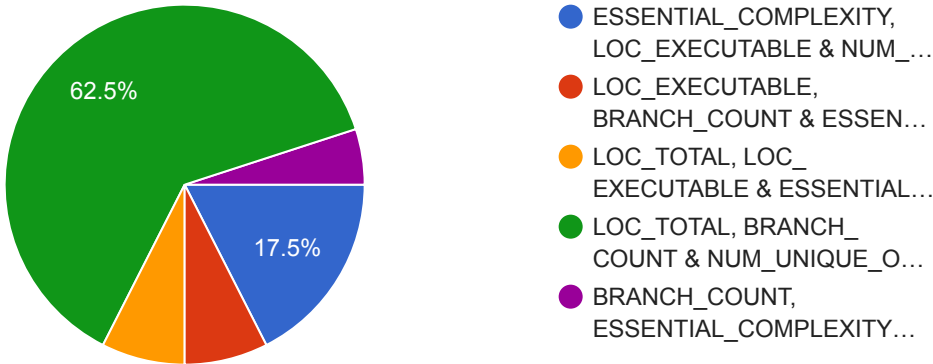
Baseado na explicação acima, você considera que controlar o LOC_TOTAL (total de linhas de código) é mais importante para evitar defeitos no código do que controlar o NUM_UNIQUE_OPERANDS (número total de operandos únicos)?

40 responses



Para finalizar o estudo, marque o conjunto de três features mais relevantes para causar defeitos no software baseado na explicação mostrada acima:

40 responses



Você pode usar este espaço para emitir qualquer opinião/sugestão sobre o estudo?

10 responses

Como não existem uma legenda para o valor base, é complicado de entender o porque de existir valores negativos.

Não gostei do termo feature. Acho que são medidas. O gráfico nos leva a dizer que o número de linhas de código aumenta as chances de causar defeito, mas discordo disso. Acho que as perguntas erradas foram feitas. É como um estudo do Zimmermann et al. Não quer dizer que se tirarmos algumas letras do teclado que normalmente são relacionadas com defeitos vamos acabar com os defeitos do software. Sugiro a reflexão.

já sabia utilizar o SHAP

A explicação local não foi muito clara.

Demorei um pouco para entender o gráfico. Apesar de no final ter achado intuitivo, acredito que dê pra deixar mais entendível olhando de primeira. Talvez usando gráficos de barra ou coordenadas paralelas. Visto que é um gráfico de múltiplas features, coordenadas paralelas pode ser interessante.

This content is neither created nor endorsed by Google. [Report Abuse](#) - [Terms of Service](#) - [Privacy Policy](#).

Google Forms



