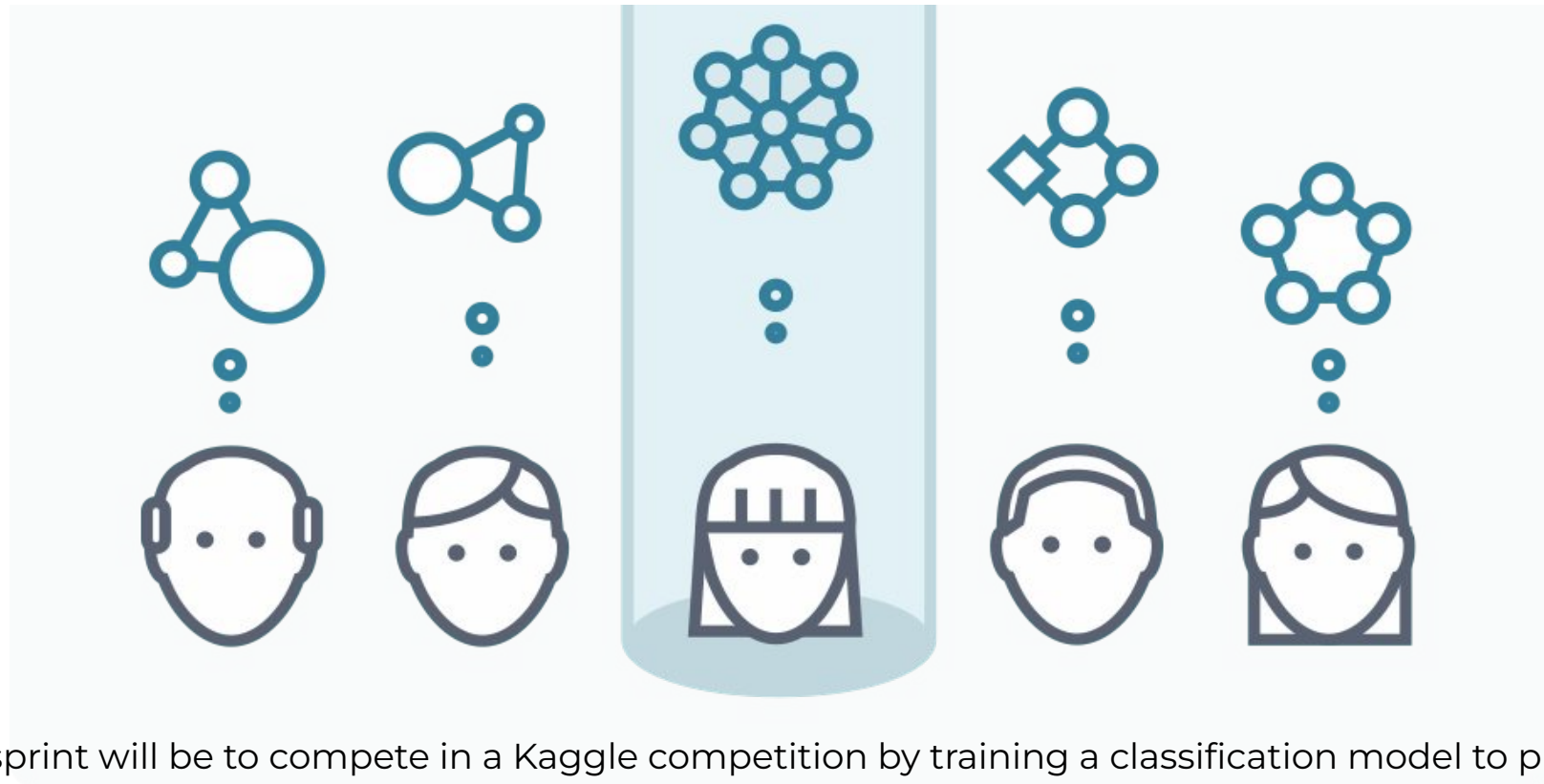


EXPLORE || DIGITAL SKILLS

Classification Hackathon

Compete in a Hackathon Challenge on Kaggle

Kaggle, according to their website, allows users to “discover and publish their own data sets, explore and build models in a web-based data-science environment; work with other data scientists; and enter competitions to solve data science challenges.”



Your challenge for this sprint will be to compete in a Kaggle competition by training a classification model to predict in which South African language a piece of text is written in.

Language Identification Challenge Summary

Enter the competition on Kaggle.

The summary below is taken directly from the competition page:



South Africa is a multicultural society that is characterised by its rich linguistic diversity. Language is an indispensable tool that can be used to deepen democracy and also contribute to the social, cultural, intellectual, economic and political life of the South African society. With such a multilingual population, it is only obvious that our systems and devices also communicate in multi-languages.

*In this challenge, you will take **text which is in any of South Africa's 11 Official languages and identify which language the text is in.***

Instructions

1. [Sign up to Kaggle](#) and **create your own personal profile** (if you have not already done so).
2. Enter the Language Identification Challenge.
3. **Create a notebook on GitHub.**
4. Develop, train and validate your classification model. **Use good practices to maintain your codebase** on GitHub.
5. **Ensure your notebook can produce a valid submission.** Submit this output to Kaggle to be placed on the Challenge Leaderboard (you can do this multiple times).
6. Forward your notebook GitHub repo link to your supervisor.



Rules

- This project is a **individual** project, therefore you are required to enter your own submission on GitHub.
- You are **not allowed to share your code, solutions, or submissions** with other individuals or teams.
- Part of your **mark will be based upon your score on the Leaderboard**.
- You will be **required to prove how you obtained a given submission result**. Those who cannot will receive a mark of 0 for the hackathon.

