**Capstone 2 : Project Ideas**

1. **Speech Recognition – Tensorflow Speech Dataset**

Use the Speech Commands Dataset to build an algorithm that understands simple spoken commands. By improving the recognition accuracy of open-sourced voice interface tools, we can improve product effectiveness and their accessibility.

1. **Sentiment Analysis - Sentiment140 dataset with 1.6 million tweets**

This is the sentiment140 dataset. It contains 1,600,000 tweets extracted using the twitter API. The tweets have been annotated (0 = negative, 4 = positive) and they can be used to detect sentiment.

1. **StackLite: Stack Overflow questions and tags**

 The dataset contains content from disparate stack exchange sites, containing a mix of both technical and non-technical questions. The problem says that we will be provided a bunch of questions. A question in Stack Overflow contains three segments Title, Description and Tags. By using the text in the title and description we should suggest the tags related to the subject of the question automatically. These tags are extremely important for the proper working of Stack Overflow.