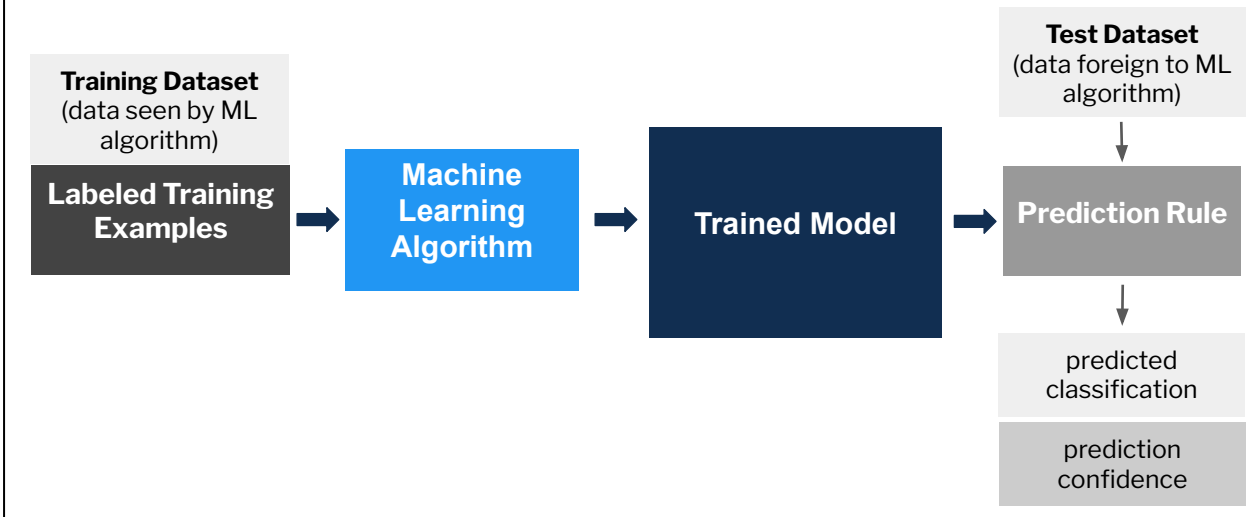
Sample Submission

Machine Learning Model



For context, a machine learning process involves multiple steps, including training data, training a model, and predicting outcomes.

Data Format

End-User License Agreements (EULAs)

- Clause Text
- Classification (Acceptable or Unacceptable)

Training Dataset (data seen by ML algorithm)	**Test Dataset** (data foreign to ML algorithm)

Classification	Clause Text
0 (Acceptable)	Node. A Node is any kind of device capable of processing data and includes, without limitation, any of the following types of computer devices: mobile/smart phone, diskless workstation, personal computer workstation, networked computer workstation, homeworker/teleworker home-based system, File Server, Print Server, e-mail server, Internet gateway device, Storage Area Network Server (SANS), terminal Servers, or portable workstation connected or connecting to a Server or network.
1 (Unacceptable)	10.3 Termination by Group Administrator. Group administrators for a Service such as COMPANY may terminate a user's access to a Service at any time. If your group administrator terminates your access, then you may no longer be able to access content that you or other users of the group have shared on a shared workspace within that Service.

My approach was to take the sample data provided for the challenge, and divide it into a training dataset and a test dataset. I used an 80-20 split.

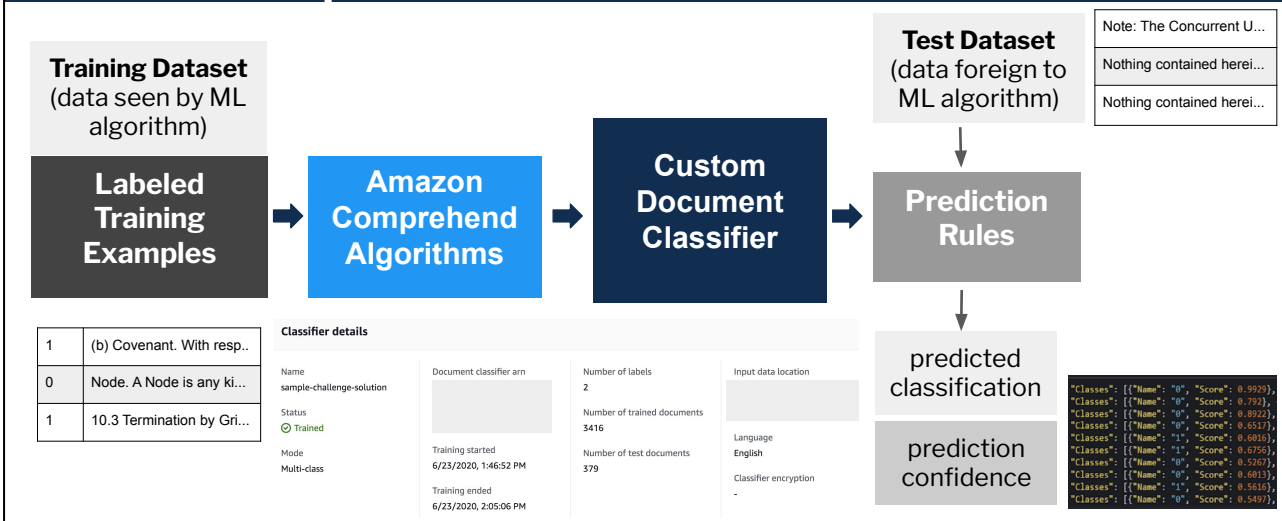
My Solution

User Interface

(Add demonstration of user interface here)

Note: this sample does not display the user interface. But you would include that here.

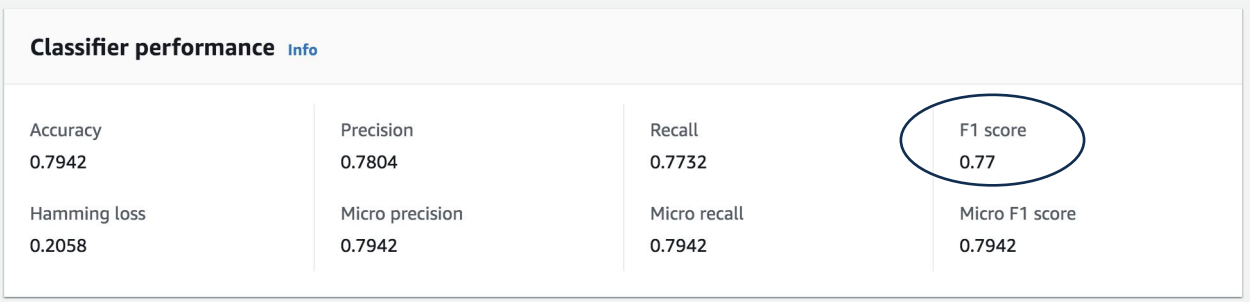
Building a Custom Classifier in AWS Comprehend



This shows the process of using AWS Comprehend to perform the machine learning.

Note: this is not an endorsement of this product.

Classifier Results



You can see the results of my classification. My overall scores were Brier Score: .25, F1 score: .77