

**Project Development Phase**  
**Model Performance Test**

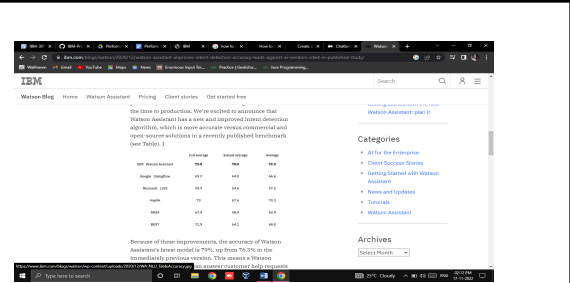
Date	18 November 2022
Team ID	PNT2022TMID31091
Project Name	AI BASED DISCOURSE FOR BANKING INDUSTRY

**Model Performance Testing:**

Project team shall fill the following information in model performance testing template.

S.No.	Parameter	Values	Screenshot
1.	Model Summary	-	<p><b>Privacy</b></p> <p>Don't store or pass any confidential or protected health information (PHI) when you interact with an AI system. This guideline applies to both standard and dedicated deployment options.</p> <p>AI systems store users' conversations or interactions in the form of logs. Those logs might be used for machine learning model improvement. Although the standard for IBM Cloud offerings is to not share any log information, it might be necessary to provide a means to opt out of this capability completely. Services such as Watson Assistant allow the customer or user to opt out of logging.</p> <p><b>Region and language support</b></p> <p>When you deploy applications that involve multiple geographies and languages, you might need to deploy the services in multiple regions by using the IBM Cloud region settings.</p> <p>AI systems must be designed and trained against various languages based on the support that is provided by the service. It is the responsibility of the application or the solution to pass the language parameters to the APIs during runtime.</p> <p><b>Performance and scalability</b></p> <p>To support a large volume of users, create a testing plan that involves load testing. You can use open source frameworks such as JMeter or third-party services such as BlazeMeter in IBM Cloud to create and run load tests.</p>

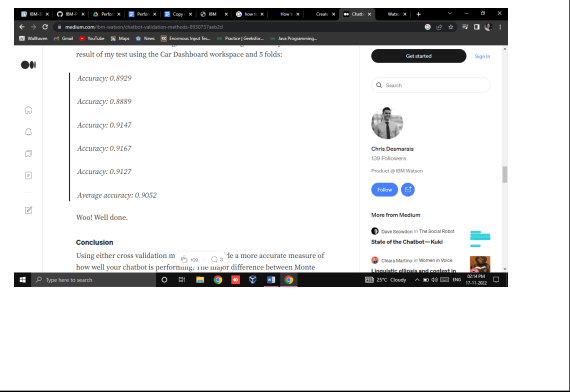
2.	Accuracy	Training Accuracy - 72.2  Validation Accuracy -0.9052
----	----------	---



IBM Watson Assistant: The time to production. We've wanted to announce that Watson Assistant has a new and improved intent detection algorithm, which is more accurate across conversational and open-source solutions in a recently published benchmark (over 700,000).

	Training Accuracy	Validation Accuracy	Average
IBM Watson Assistant	72.2	72.2	72.2
Google Dialogflow	67.2	68.2	67.7
Amazon Lex	58.2	59.2	58.7
Microsoft LUIS	70	67.2	68.6
Dialogflow ES	67.2	68.2	67.7

Because of these improvements, the accuracy of Watson Assistant's intent results is 72.2%, up from 70.5% in the benchmark published recently. This means a Watson Assistant-powered solution will be more accurate.



result of my test using the Car Dashboard workspace and 5 folds:

- Accuracy: 0.8929
- Accuracy: 0.8889
- Accuracy: 0.9147
- Accuracy: 0.9167
- Accuracy: 0.9127
- Average accuracy: 0.9052

Wow! Well done.

**Conclusion**

Using either cross validation or  $\frac{1}{n}$  is a more accurate measure of how well your classifier is performing. Use a small difference between training and validation accuracy.