ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

Machine Learning to Predict Human Override

Mayuresh Patil

Digital Product Manager/ Arch, Digital product Development Fenchurch St, Billingsgate, UK

ABSTRACT: We have built around the idea that we would be able to build a machine in few days to find patterns in the historical data of overridden results and for future rules evaluations, this machine would be able to recommend an override reason. In Machine Learning, data is key.. The key was to find the right type of data, and in the right amount, to avoid statistical sampling errors. Since the ramp up time was about three days (with another half for presenting a demo), we needed something that could provide good predictive results, without diving deep into data science technologies. After exploring the technology space, we found that libraries lApache MXNet and Tensorflow, and languages like R, were the popular choice for deep learning enthusiasts. Now while these are awesome technologies, because of the time constraint, and the lack of an expert data scientist on the team, we kept looking for more options. This paper attempts to provide an insight in AWS Machine Learning API provided predictions with quite good accuracy to override the human intervention in operating machines.

KEYWORDS: Machine learning, artificial intelligence, deep learning, neural network, chatbot.

I. INTRODUCTION

We selected AWS machine learning considering its benefits:

- Provides complex deep learning capabilities via a simplified API in the cloud.
- Reduces the barrier to entry in machine learning for businesses not invested heavily into data science
- Democratizes access to machine learning for developers not strong on statistical understanding.

The AWS ML API places strong emphasis on the understanding of the business domain, thus helping businesses using their own existing domain experts to help drive machine learning for them.

Data Attributes and Result Classification

We contacted our client teams to understand the structure of the data, and the key attributes and factors that workers would traditionally key-off when making a human decision to override the system result.

No override was exercised in about 80% of the scenarios. In the remaining 20%, when the workers used override, the following reasons were being used by them to override the system results:

1. Admin
2. Appeal Pending
3. Citizenship
4. External System
5. Internal System
6. Incorrect Config
7. Irregular Tax Filing

We then started preparing datasets using this information. In the first run, we came up with a training dataset of 2007 records with 20 significant attributes (total 22).

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

II. RELATED WORK

Zhao Jianqiang et al.[1] proposed that emotion-aware mobile applications have been increasing due to their smart features and user acceptability. To realize such an application, an emotion recognition system should be in real time and highly accurate. In this paper, emotion recognition with high performance for mobile applications is proposed. In the proposed system, facial video is captured by an embedded camera of a smart phone. Some representative frames are extracted from the video, and a face detection module is applied to extract the face regions in the frames. The dominant bins are then fed into a Gaussian mixture modelbased classifier to classify the emotion. Experimental results show that the proposed system achieves high recognition accuracy in a reasonable time. The merits of this model include enhanced recognition features, easier implementation and quick response. The demerits of this model include less speed, unsuitable for large datasets and increased overhead.

M. Shamin Hossain et al.[2] proposed that the Twitter sentiment analysis technology provides the methods to survey public emotion about the events or products related to them. Most of the current researches are focusing on obtaining sentiment features by analyzing lexical and syntactic features. These features are expressed explicitly through sentiment words, emoticons, exclamation marks, and so on. In this paper, a word embeddings method is used which is obtained by unsupervised learning based on large twitter corpora, this method using latent contextual semantic relationships and cooccurrence statistical characteristics between words in tweets. The project had made provisions to analyze real time data. The implementation of the model is practical and more realistic. The project also classifies emotions based on emojis. The project involves complex interpretations, increased overhead. The response model is not direct and involves collective responses. Jian Guo et al.[3] proposed that emotion recognition has a key role in affective computing. A compound facial emotion includes dominant and complementary emotions (e.g., happily-disgusted and sadly-fearful), which is more detailed than the seven classical facial emotions (e.g., happy, disgust, and soon). To address these problems, the iCV-MEFED dataset is released, which includes 50 classes of compound emotions and labels assessed by psychologists. The task is challenging due to high similarities of compound facial emotions from different categories. However, the proposed data set can help to pave the way for further research on compound facial emotion recognition. The model produces accurate feature extraction results. The processing and computing speed were high. There was no real time analysis of data involved in the project. More stable algorithm is needed since the model may be prone to change. Mondher Bouazizi et al.[4] proposed that with the rapid growth of online social media content, and the impact these have made on people’s behavior, many researchers have been interested in studying these media platforms. A major part of their work focused on sentiment analysis and opinion mining. These refer to the automatic identification of opinions of people toward specific topics by analyzing their posts and publications. The dataset was manually labeled and the results of the automatic analysis were checked against the human annotation. The experiments show the feasibility of this task and reach an F1 score equal to 45.9%. The model classifies wide range of data. It provides enhanced real time sentiment analysis. Variety of data has been analyzed. The project module has hierarchical dependency of different algorithms making it complex and less susceptible to maintenance. Guixian Xu et al.[5] proposed that with the rapid development of Internet technology and social networks, a large number of comment texts are generated on the Web. In the era of big data, mining the emotional tendency of comments through artificial intelligence technology is helpful for the timely understanding of network public opinion. The technology of sentiment analysis is a part of artificial intelligence, and its research is very meaningful for obtaining the sentiment trend of the comments. The essence of sentiment analysis is the text classification task, and different words have different contributions to classification. In the current sentiment analysis studies, distributed word representation is mostly used. The experimental results show that the proposed sentiment analysis method has higher precision, recall, and F1 score. The method is proved to be effective with high accuracy on comments. Xin Kang et al.[6] proposed that understanding people’s emotions through natural language is a challenging task for intelligent systems based on Internet of Things (IoT). The major difficulty is caused by the lack of basic knowledge in emotion expressions with respect to a variety of real world contexts. In this paper, a Bayesian inference method is proposed to explore the latent semantic dimensions as contextual information in natural language and to learn the knowledge of emotion expressions based on these semantic dimensions. The Bayesian inference results enable us to visualize the connection between words and emotions with respect to different semantic dimensions. And by further incorporating a corpus-level hierarchy in the document emotion distribution assumption, we could balance the document emotion recognition results and achieve even better word and document emotion predictions. The model is simple to deploy and real time data analysis was made possible.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

The model can only synthesize texts. Wanliang Tan et al.[7] proposed that sentiment analysis of product reviews, an application problem, has recently become very popular in text mining and computational linguistics research. Here, the correlation between the Amazon product reviews and the rating of the products given by the customers need to be studied. Both traditional machine learning algorithms including Naive Bayes analysis, Support Vector Machines, K-nearest neighbor method and deep neural networks such as Recurrent Neural Network (RNN), Recurrent Neural Network (RNN) are used. By comparing these results, better understanding of these algorithms can be obtained. They could also act as a supplement to other fraud scoring detection methods. The project involves feedback oriented analysis, making it more realistic and best fit for practical applications. Data contains only customer reviews, because of which model is trained in a unidirectional path. Tzuu-Heseng S.Li et al.[8] proposed that Facial expression recognition (FER) is a significant task for the machines to understand the emotional changes in human beings. The result of final recognition is calculated using softmax classification. Fine-tuning is effective to FER tasks with a well pre-trained model if sufficient samples cannot be collected. The model contains well-constructed algorithm and is highly stable. It is really slow in computation wise and requires large processing power.

III. PROPOSED SOLUTION

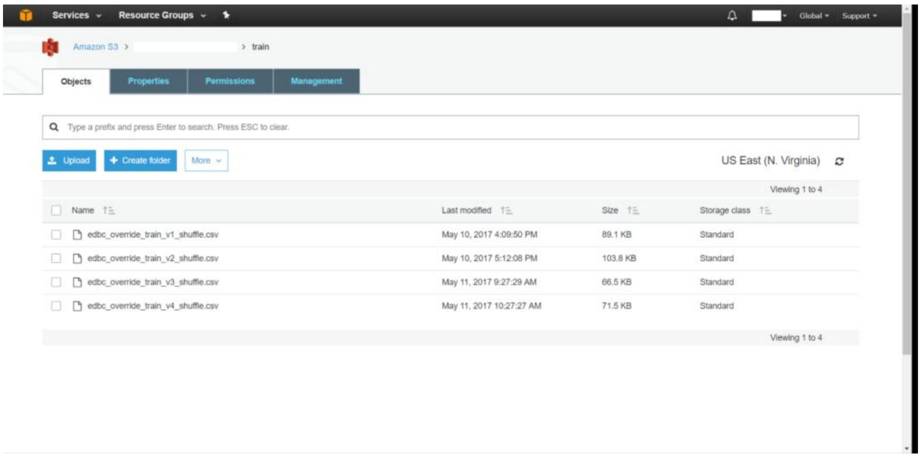
Data Structure and StorageThe training dataset for AWS needs to be in a CSV format. We need to tell the AWS ML API where to get the data from. For this exercise, since our attributes were all Non-Personal-Information (NPI), we decided to go ahead with S3, Amazon Simple Storage Service to manage the training data.

We shuffled the original dataset before saving it to S3 to ensure randomness.

```
$ tail -n+2 edbc_override_train_v1.csv | shuf -o edbc_override_train_v1_shuffle.csv
$ head -1 edbc_override_train_v1.csv | cat - edbc_override_train_v1_shuffle.csv > temp && mv temp

edbc_override_train_v1_shuffle.csv
```

AWS S3 Folders and Training Files



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

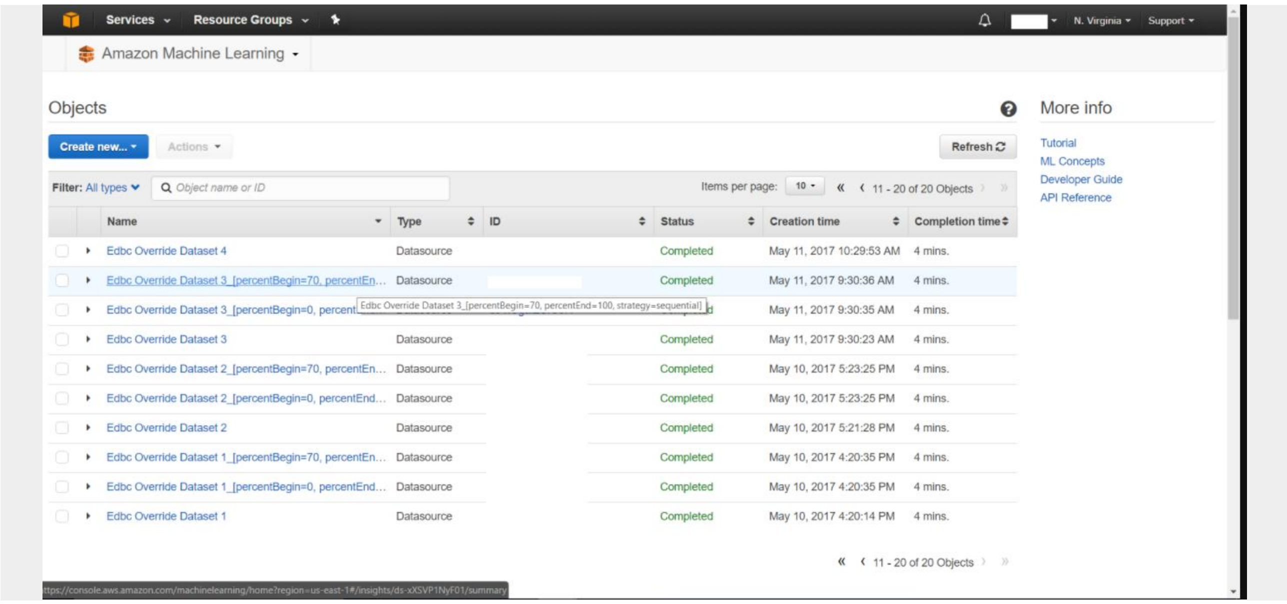
Vol. 7, Issue 6, June 2019

AWS Modeling

We then created a model using the AWS Console's guided wizard with default values selected for most options.

To create the model, AWS uses 70% of the training data to prepare the model, then uses the remaining 30% of the data to verify its accuracy and come up with an accuracy score (these figures are all adjustable, this blog only discusses defaults).

AWS automatically knew this was a multi-class model (since there could be 8 classes of overrides – 7 listed above with the 8th being no override).



AWS ML Dashboard with Strategy for Splitting

To mention the output is a set of prediction scores for multiclass classification algorithm. The indication of the score is driven by the model and its certainty for the stated observation. (e.g. a given record in our csv) belongs to each of the classes.

For the first dataset, the Service created a multi-class model for us with an F1 score of 0.883.

Model Accuracy (Numeric): In Amazon ML, the macro-average F1 score is used to evaluate the predictive accuracy of a multiclass metric.

The usual range is from 0 to 1 and it has co-relation as larger the value better the accuracy and value towards 0 indicates the dip in accuracy. F1 score which is usually indicated as binary classification which considers both recall value and metrics precision.

For calculations usually the unweighted average of the F1 score for overall for all the classes. consider the frequency of occurrence of the classes in the evaluation dataset. A larger value indicates better predictive accuracy.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

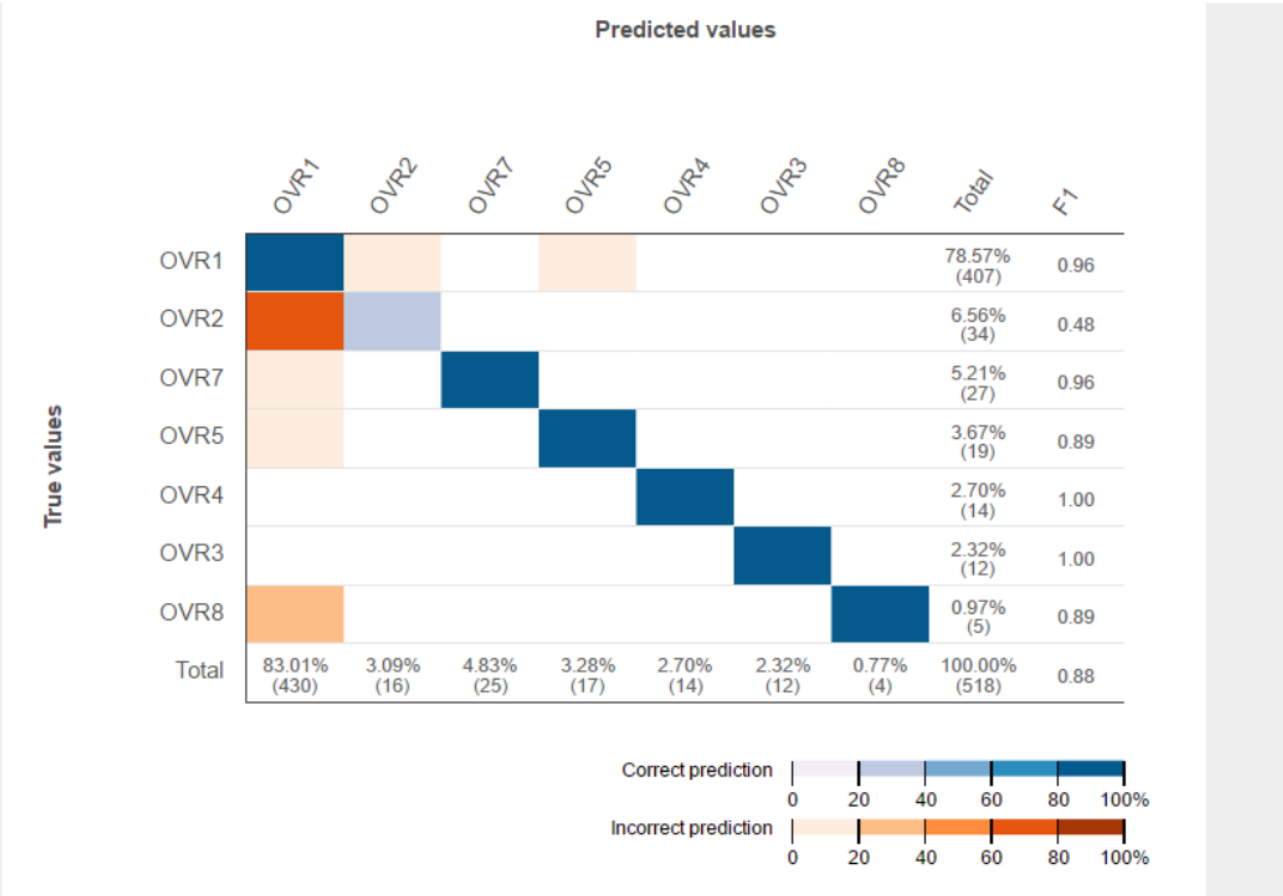
Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

Model Accuracy (Visual):

Amazon ML provides a confusion matrix for various predictive models to give a view of classified accuracy of multiclass. The matrix illustrates either number or percentage of all kind of predictions by each class and the comparison will happen on observations, predictions.

Here is a matrix showing the true versus predicted values for the model.



Confusion Matrix for Model 1

The blue bars indicate a good prediction accuracy, and ideally, we would want to have all diagonal bars as blue. But that is terribly difficult, and perhaps unwise to attempt.

For us, we wanted our error threshold to be near 20%, so our model was not yet mature. As seen in the matrix, it was getting 2 classifications, OVR2 and OVR8, wrong, over 60% and over 40%, of the time.

For quick reference, here is how the OVR#s map back to our override reasons.

"OVR1" : "No Override",
"OVR2" : "Admin",

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

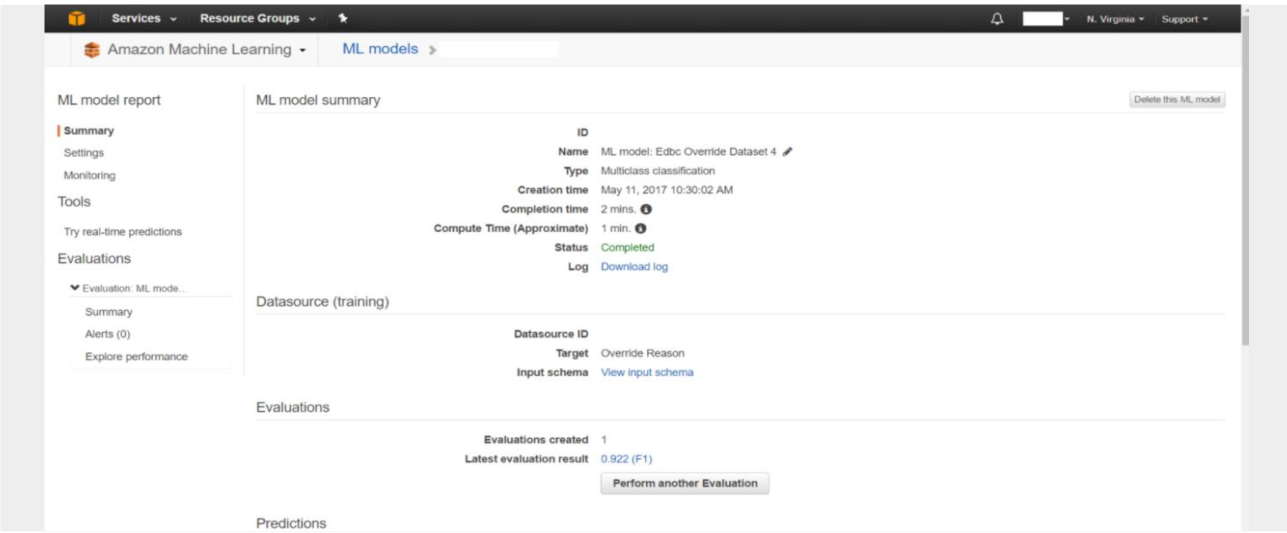
(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

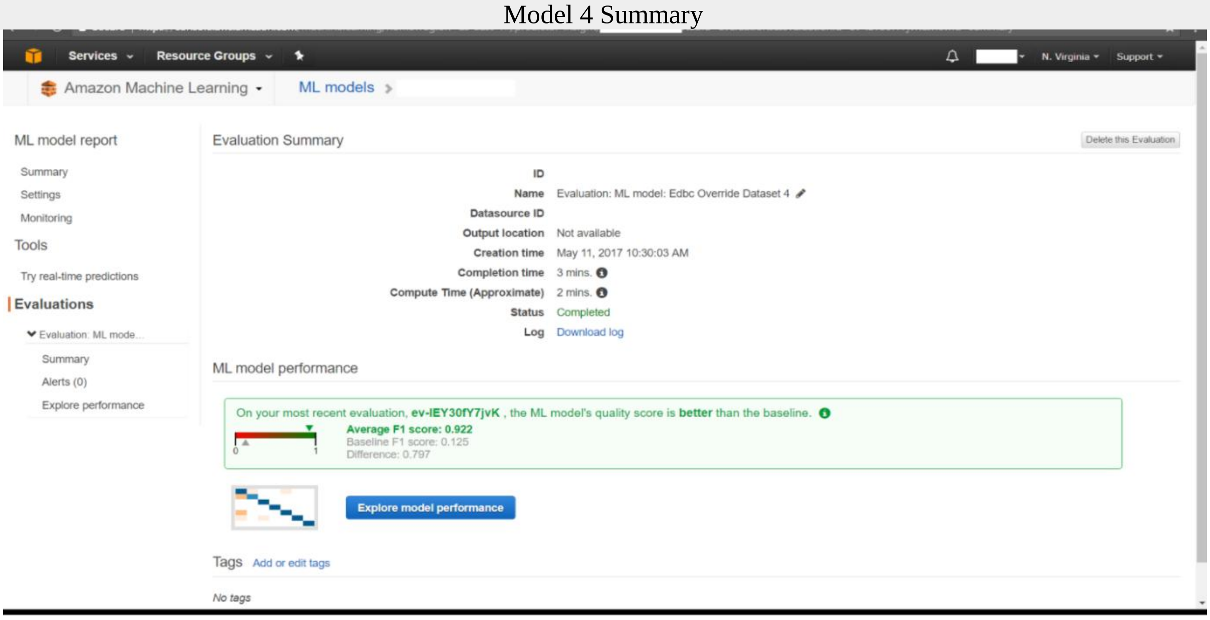
Vol. 7, Issue 6, June 2019

"OVR3" : "Appeal Pending", "OVR4" : "Citizenship", "OVR5" : "External System", "OVR6" : "Internal System", "OVR7" : "Incorrect Config", "OVR8" : "Irregular Tax Filing"

We iterated over the preparation of training dataset with better data, and revised significant attributes, till on the 4th attempt we landed with an F1 score of 0.922, and errors in predictions contained to less than 20% for all but two classifications (which were also good at errors in prediction less than 40% of the time).



Model 4 Summary



Model 4 Performance

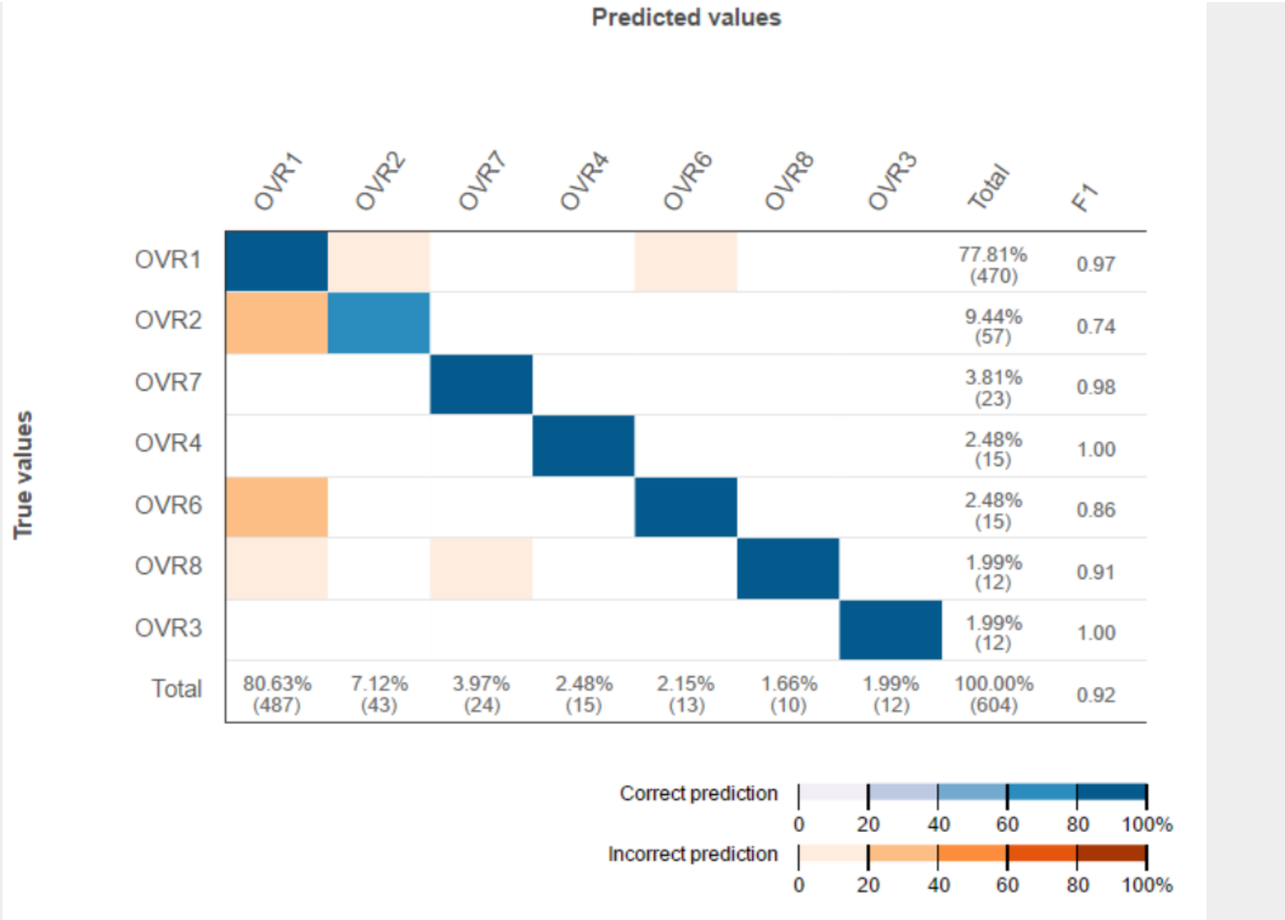
ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019



Model 4 Confusion Matrix

Demo

Once we were satisfied with our modeling accuracy, for the demo, we used a simple web-application where users could enter test data (or select some pre-loaded test data that was displayed on screen).

AWS Services Used

AWS S3 – Simplified Storage Service API – Stores Training Data
AWS ML – Machine Learning API – Recommends Classification (i.e. Override Reason)
AWS IAM – Authenticates Developer Access to Amazon Web Services APIs

In our web-app, the front end (JSP) was making a simple rest call with JSON-ified form data to a RESTful service in the backend.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

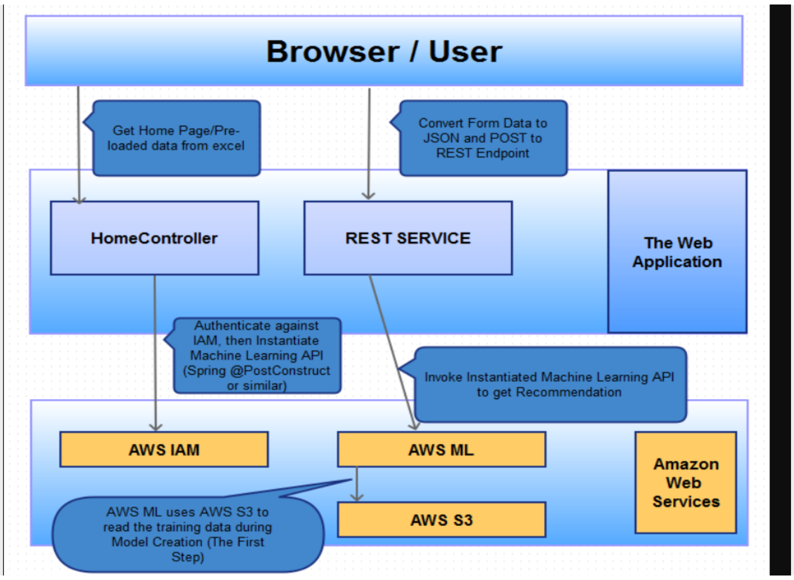
International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

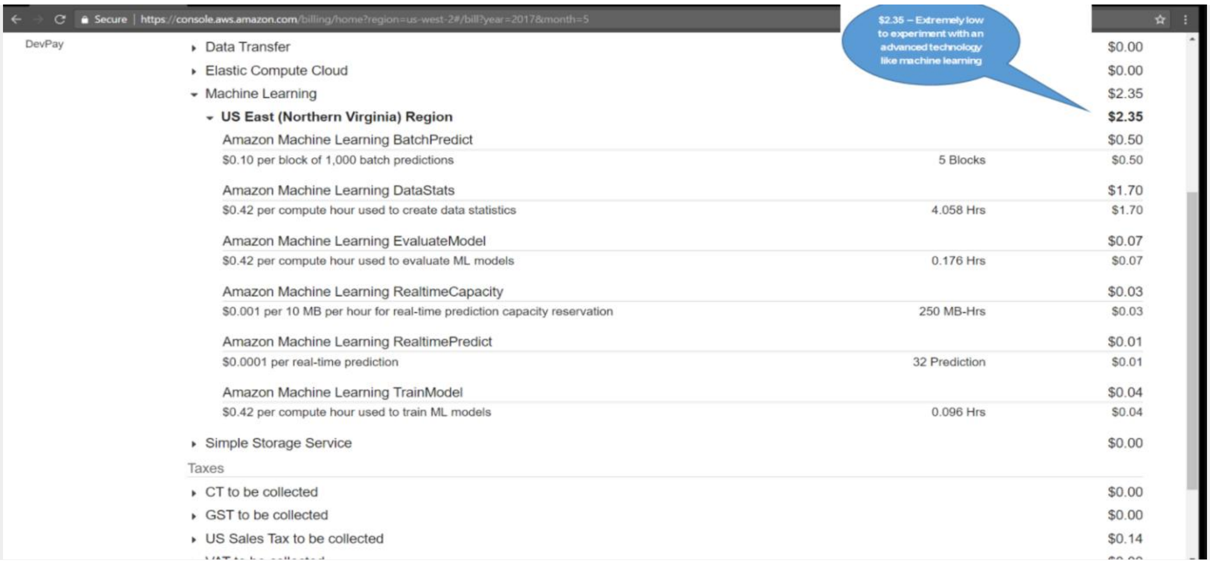
This RESTful service took the form data, converted it to a format described by the AWS ML API,and invoked the AWS ML API using the authentication credentials setup in the AWS IAM. The results were then converted back to JSON and returned to the front-end for display.



Demo System Interaction Diagram

Costs

AWS charges users based on the size of your model. For a period of two days, here are the charges we incurred.



Data Transfer		$0.00
Elastic Compute Cloud		$0.00
Machine Learning		$2.35
US East (Northern Virginia) Region		$2.35
Amazon Machine Learning BatchPredict		$0.50
$0.10 per block of 1,000 batch predictions	5 Blocks	$0.50
Amazon Machine Learning DataStats		$1.70
$0.42 per compute hour used to create data statistics	4.058 Hrs	$1.70
Amazon Machine Learning EvaluateModel		$0.07
$0.42 per compute hour used to evaluate ML models	0.176 Hrs	$0.07
Amazon Machine Learning RealtimeCapacity		$0.03
$0.001 per 10 MB per hour for real-time prediction capacity reservation	250 MB-Hrs	$0.03
Amazon Machine Learning RealtimePredict		$0.01
$0.0001 per real-time prediction	32 Prediction	$0.01
Amazon Machine Learning TrainModel		$0.04
$0.42 per compute hour used to train ML models	0.096 Hrs	$0.04
Simple Storage Service		$0.00
Taxes		
CT to be collected		$0.00
GST to be collected		$0.00
US Sales Tax to be collected		$0.14

AWS Machine Learning Demo Cost Structure

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

IV. CONCLUSION AND FUTURE WORK

The results showed that the In four days, our team successfully created and demonstrated a machine learning model to find patterns in case worker outcomes and recommend actions on future cases based on historical rules determination and data attributes. A significant chunk of our time was devoted to coming up with the right modeling attributes and appropriately tuning the model. With this, the AWS Machine Learning API provided predictions with quite good accuracy. Finally, if you see the costing above, the AWS Machine Learning API enables individuals and businesses to play around with machine learning themselves, without much help, at very reasonable costs, to see if machine learning would be beneficial to them in some way.

REFERENCES

[1] Zhao Jianqiang, Gui Xiaolin and Zhang Xuejun- School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China, "Deep Convolution Neural Network on Twitter Sentiment Analysis", IEEE Conference on Development of personal Humanoids for better human understanding (2017)

[2] M. Shamin Hossain- Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia, Ghulam Muhammad - Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia, "An Emotion Recognition System for Mobile Applications", IEEE Special Section of emotion-aware mobile computing(2017)

[3] Jian Guo, Zhen Lei, Jun Wan, Eglis Avots – Dept. of Electrical and Electronic
Engineering, Hasan Kalyponeu University, Turkey, "Dominant and Complementary Emotion Recognition from Still Images of Faces", IEEE Special Section on Visual Surveillance And Biometrics (2018)

[4] Mondher Bouazizi and Jomoaki Ohtsuki – Graduate School of Science and Technology, Keio University, Yokohoma, Japan, "Multi-Class Sentiment Analysis in Twitter", IEEE special Section on Emerging Trends, Issues and Challenges of Array signals (2018)

[5] Guixian Xu, Yueting Meng, Xiaoyu Qiu, Ziheng Yu and Xu Wu – College of Information Engineering, Minzu University of China, Beijing, China, "Sentiment Analysis of Comment Texts Based on BiLSTM", IEEE Special Section on Artificial Intelligence and Cognitive Computing for Communication and Network (2019)

[6] Xin Kang, Fuji Ren, Yunong Wu, IEEE Members, "Exploring Latent Semantic Information for Textual Emotion Recognition in Blog Articles", IEEE Special conference on Automation (2018)

[7] Wanliang Tan, Xinyu Wang, Xinyu Xu – Department of Computer Engineering, Stanford University, USA, "Sentiment Analysis for Amazon Reviews", International Conference on Human and AI interaction (2018)

[8] Tzuu-Heseng S.Li, Ping-Huan Kuo, Ting-Nan Tsai and Po-Chien Luan- Department of Electrical Engineering, National Cheng Kung University, Pingtung, Taiwan, "CNN and LSTM Based Facial Expression Analysis Model for a Humanoid Robot", IEEE conference on Human and AI interaction (2019)

[9] D. Yangm Abeer Alsadoon, P.W.C. Prasad, School of Computing and Mathematics, Charles Sturt University, Sydney, Australia, "An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment", International Conference on Smart Computing and Communications (2017)

[10] Carlos Busso, Zhigang Deng, Serdar Yildirim, Viterbi School of Engineering, University of Southern California, Los Angeles, Analysis of "Emotion Recognition using Facial Expressions, Speech and Multimodal