Technical Disclosure Commons

May 14, 2019

FABLER-MACHINE LEARNING BASED CONTEXT AWARENESS IDENTIFICATION PLATFORM

HP INC

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

FABLER- Machine Learning Based Context Awareness Identification Platform

Abstract

The explosive growth of users of Smart Phone, Web Connected devices and Web has changed how content was absorbed and presents tremendous opportunity for Data acquisition, Data Mining, Data Rendering, Data Classification and Visualization. A leader in the Printing and Personal System domain we have unique opportunity to direct this data to Cloud Print enabled printers and various Web Connected Devices for providing various solutions.

This paper aims at defining a machine learning based context awareness identification platform for data acquisition, extracting context from unstructured data and helping various data driven solution e.g. enhancing user experience of digital document reading by adding context awareness using machine learning on top of existing text to audio translation solutions available in the market, Real time customer support system for cost reduction, reading unknown language and sign board, Enhancing Corporate Training Experience and Increasing Employee Engagement and printable content Identification etc.

Problem statement

According to a recent study, Worldwide Big Data market revenues for software and services are projected to increase from \$42B in 2018 to \$103B in 2027, attaining a Compound Annual Growth Rate (CAGR) of 10.48%. The amount of data we produce every day is truly mind-boggling. There are 2.5 quintillion bytes of data created each day at our current pace, but that pace is only accelerating with the growth of the Internet of Things (IoT). With this exponential data growth identifying relevant information and using that information in proper context is big challenge. Below are few scenarios where context identification can add tremendous value for providing various solution-

1. **Digital ebook reader:** Current digital ebook reader or document reader just do text to audio conversion, they lack in provding context awareness and emotions, voice pitch, modulation, male/female voice switchover according to scenario, region, country and locale specific tone and voice support. Readers lack cinematic experience which usually they feel while watching a movie.

2. **Call center Support:** Call centres are the front line of all big organisations, providing a vital link between businesses and customers. They are also the place that can make or break the brand experience. With the rapid growth of data and customers, currently Call centres facing two major problems. First, How to seamlessly give customers the right information they need at the right time by offering self-service options, eliminating the need for a call to customer service. Second, How to give customer service representatives more information to help them handle the complicated issues that self-service cannot resolve.

3. **Unknown language and sign board Reader:** Suppose you visiting a new country and you don't understand language for that particular country. It will be very difficult for you to understand documents, sign board and instruction written in the language which you don't understand.

4. **Corporate Training and Employee Engagement-** Other scenario is Corporate training documents, using solution proposed by us, it will make training more engaging, effective and closer to reality when it's about getting feedback and evolving the training modules and ultimately increasing the employee engagement index of the organization.
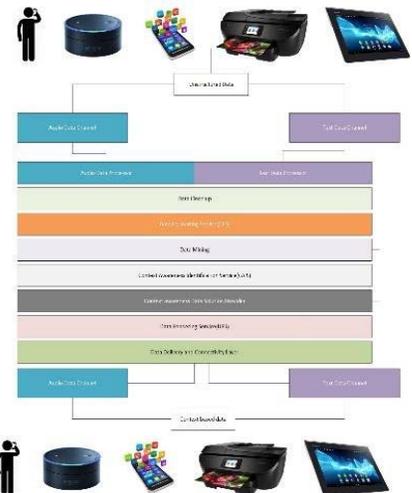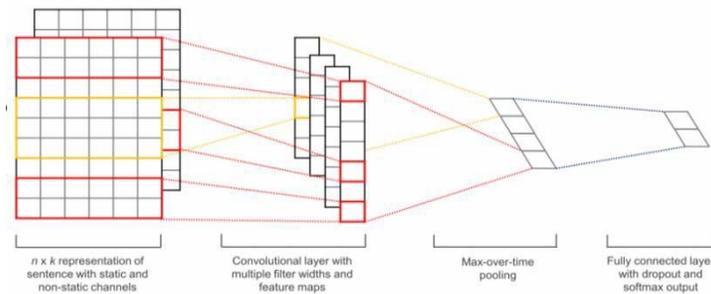
These are some of the diverse set of the problems which are very critical to the customers. With the help of Machine Learning based Context Awareness Identification Platform can help customers in solving them.

1

Our solution

The solution is to build a platform which can take any kind of data as an input and provide the context of that data as an output. Using the Fabler Flatform we are building **Digital eBook Document Reader** by using which Digital Documents will be parsed and processed by our machine learning algorithm and will extract/identify the context of the content such as whether content is a conversation or narration, participating characters and their personality traits, plot and emotions involved within. Based on that, audience will get the experience of listening the document as close to reality as possible, like we get the experience in cinematic audio. In our solution we have applied **Convolutional Neural Network (CNNs)** to problems in **Natural Language Processing (NLP)** for Sentence Classification and gotten some interesting results. As shown in below sequence **diagram-1**-
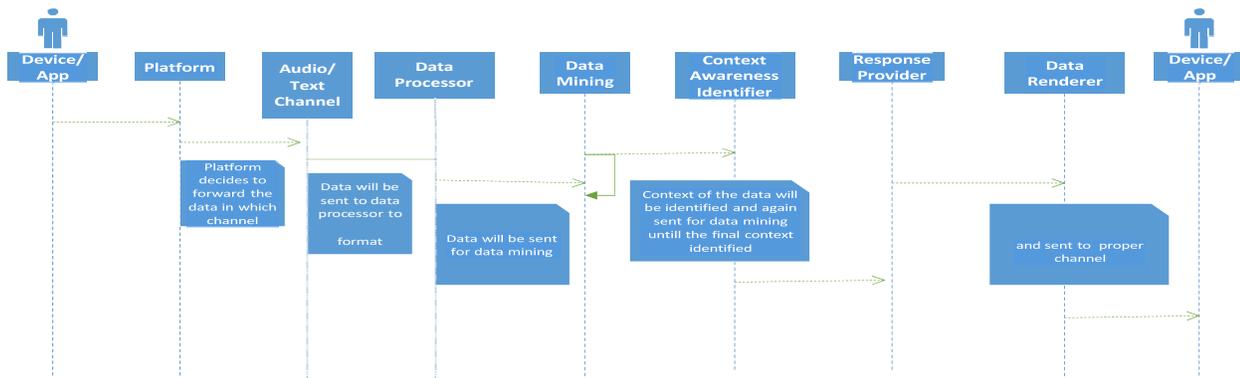
1. User reads an ebook or makes a support call from a web connected device like mobile,tablet,pc etc.. While reading ebook he selects to translate the ebook with context, A REST based web services call will be made to our Context platform with request multi-part body as the PDF/ DOC file and channel as text-channel, Auth token (Security Channel) in Request header.

2. The platform will Read the incoming Request and Forward the Request to the Text-channel, The Text channel/Audio channel are the data streams. The data will be Collected by the **DATA processor (DP)** , The Data Processor will use a Machine Learning Algorithm to Clean unwanted data **(Data Filtering Engine)** and send the Usefull data for Data Mining and classification.

3. A **Context Awarenes Identifier Service (CAIS**) will use **Natural Language Processing (NLP)**, **Convolutional Neural Network(CNN)** and Machine Learning Algorithm to identify the context of the data.

4. The Network model of CNN ,NLP and ML will look like the following model shown in **diagram-2** and **diagram-3**-

n x k representation of sentence with static and non-static channels   *Convolutional layer with multiple filter widths and feature maps*   *Max-over-time pooling*   *Fully connected layer with dropout and softmax output*

5. The first layers embeds words into low-dimensional vectors. The next layer performs convolutions over the embedded word vectors using multiple filter sizes. For example, sliding over 3, 4 or 5 words at a time. Next, we max-pool the result of the convolutional layer into a long feature vector, add dropout regularization, and classify the result using a softmax layer. Pixels are made of embedding vectors of each word in a sentence· Convolutions are performed based on word-level· Classify each sentence based on the class vector (In our Digital eBook Reader use case It's Male_Gender, Female_Gender,

2

Neutral_Gender, Sad_Emotion, Happy_Emotion, Neutral_Emotion, Narration_Conversation, Neutral_Conversation, High_Voice_Pitch, Low_Voice_Pitch, Neutral_Voice_Pitch). So, the two methods will be followed here 1) CNN using 1D convolution and pooling. 2) CNN using 2D convolution and pooling.

6. Once the context is identified the data will be sent to **Response Provider Service**(RPS), Here the context data will be mapped with characters with voice modulation and pitch.The **Data Renderring Service**(DRS) will receive input from RPS as context data,characters,voice,pitch as parameters and renders accordingly and generates an audio output file which will be again sent to the audio response channel stream. The data from stream will be returned back to the user.

Evidence the solution works

We have created a POC for Digital eBook Reader business use case for context awareness identification. Our initial results are very encouraging in effectively addressing the original problem. We have successfully able to extract the context from the pdf document with 96% accuracy as shown in the **diagram-4**. This is the result of the small set of training data set.

```
Correctly Classified Instances          29              96.6667 %
Incorrectly Classified Instances        1               3.3333 %
Kappa statistic                         0.9628
Mean absolute error                     0.0061
Root mean squared error                 0.0775
Relative absolute error                 3.626  %
Root relative squared error             26.5689 %
Total Number of Instances               30

=== Detailed Accuracy By Class ===

              TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
              1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Male_Gender
              1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Female_Gender
              1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Neutral_Gender
              1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Sad_Emotion
              1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Happy_Emotion
              1.000    0.038    0.800      1.000   0.889      0.877  1.000     1.000     Neutral_Emotion
              1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Narration_Conversation
              1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Neutral_Conversation
              0.500    0.000    1.000      0.500   0.667      0.695  0.661     0.548     High_Voice_Pitch
              1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Low_Voice_Pitch
              1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Neutral_Voice_Pitch
Weighted Avg. 0.967    0.005    0.973      0.967   0.963      0.963  0.977     0.970
```

Competitive approaches

There are no prior solutions available for addressing these kinds of problems. Current digital ebook reader (Google Ebooks and Amazon Kindle) or document reader just do text to audio conversion only. Traditional Call centres Support system having a huge cost factor involved. Providing Context Identification platform will be huge differentiating factors and can add lot of value to customers.

Current status

We already have POC ready for our solution to convert the pdf document into the speech (Text to Speech Conversion Engine). We also have machine learning based solution ready to extract context from the pdf document. Context Identification REST interface is available where user can provide pdf file as input and get context of the document as output. We have already tested this for various kinds of data. During our initial phase of solution validation we have received very good results and accuracy.

Next steps

We are planning to finish first Digital eBook Reader solution end to end and open the interfaces for other business use cases. We are also in process of refining regression model and broadening training data for final integration. Along with this we also looking for indentifying optimal cost function and re-structuring of framework, to create scope for introducing value added function on top of this solution. We are also in the process of filing few IP's on the Context awareness Identification Method for different data types to protect our idea. We will be building other use cases (e.g. Call center Support, Unknown language and sign board Reader, Corporate Training and Employee Engagement etc.) as well using the Context awareness Identification platform. Context awareness identification platform can also be used for business use cases such as businesses gaining understanding about consumer reactions to a product or detecting hateful speech in online comments by extracting the emotional content in text and identifying printable content from the web.

References

[1] CNN : https://www.researchgate.net/publication/285164623_An_Introduction_to_Convolutional_Neural_Networks
[2] Natural Language Processing : http://www.academia.edu/Documents/in/Natural_Language_Processing
[3] Logistic Regression Model(Machine Learning) : https://en.wikipedia.org/wiki/Logistic_regression
[4] Supervised learning : https://en.wikipedia.org/wiki/Supervised_learning
[5] CN N/W : https://www.sciencedirect.com/science/article/pii/S1877050918306197

3

Disclosed by
Devraj
Singh, HP
Inc.