# Named Entity Recognition for Automotive Domain using Fine-Tuned Large Language Model


Name: Souvik Ghosh

Organization: Predii

Subject: Assignment Round for ML/NLP Engineer Internship

Date: 08-03-2024

## Table of Contents

## Objective

Fine tune a large language model (LLM) for performing named entity recognition (NER) task on an automotive dataset. NER task involves processing unstructured text data to extract useful information/entities.

This assignment is broken down in three tasks:

First task involves analysing the data and identifying what are some automotive entities that can be extracted from this data. We are interested in entities related to automotive domain. Some examples could be component, failure issue, vehicle model, corrective action etc.

Second task is to use an open source LLM and write the prompt to extract the automotive domain entities from given dataset. We have used Llama2-7b and Flan-T5 in our project. The Llama2-7b LLM has been trained using Zero Shot Learning Technique and Flan-T5 was trained using both Few Shot Learning and Zero Shot Learning technique but responses of the model on Few-Shot Learning performed better.

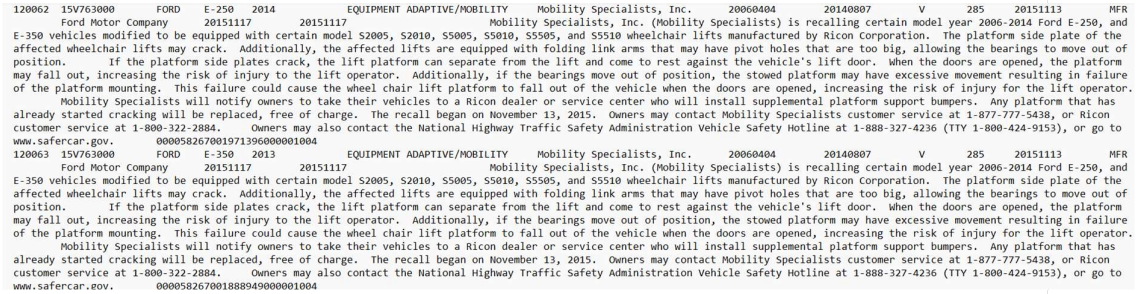
Final task is to fine tune the selected LLM on a subset of provided dataset. We have chosen Llama2-7b for Fine Tuning. The model was fine-tuned against held out dataset obtained by pre-processing the Few-Shot Learning model response of Flan-T5.

# Introduction

- Named Entity Recognition in Natural Language Processing that focusses on identifying Named Entities based on a particular domain from textual context.

Here, we focus on Fine Tuning Llama2 Large Language Model (LLM) to extract Named Entities from automotive domain. Examples of Named Entities from automotive domain such as: component, failure issue, vehicle model, corrective action and others.

- Description of the dataset: NHTSA Recall dataset (FLAT_RCL.txt) contains 256066 rows. We use data from the columns: {DEFECT SUMMARY, CONSEQUENCE SUMMARY, CORRECTIVE SUMMARY, RECALL NOTES}



120062 15V763000 FORD E-250 2014 EQUIPMENT ADAPTIVE/MOBILITY Mobility Specialists, Inc. 20060404 20140807 V 285 20151113 MFR
Ford Motor Company 20151117 20151117 Mobility Specialists, Inc. (Mobility Specialists) is recalling certain model year 2006-2014 Ford E-250, and E-350 vehicles modified to be equipped with certain model S2005, S2010, S5005, S5010, S5505, and S5510 wheelchair lifts manufactured by Ricon Corporation. The platform side plate of the affected wheelchair lifts may crack. Additionally, the affected lifts are equipped with folding link arms that may have pivot holes that are too big, allowing the bearings to move out of position. If the platform side plates crack, the lift platform can separate from the lift and come to rest against the vehicle's lift door. When the doors are opened, the platform may fall out, increasing the risk of injury to the lift operator. Additionally, if the bearings move out of position, the stowed platform may have excessive movement resulting in failure of the platform mounting. This failure could cause the wheel chair lift platform to fall out of the vehicle when the doors are opened, increasing the risk of injury for the lift operator.
Mobility Specialists will notify owners to take their vehicles to a Ricon dealer or service center who will install supplemental platform support bumpers. Any platform that has already started cracking will be replaced, free of charge. The recall began on November 13, 2015. Owners may contact Mobility Specialists customer service at 1-877-777-5438, or Ricon customer service at 1-800-322-2884. Owners may also contact the National Highway Traffic Safety Administration Vehicle Safety Hotline at 1-888-327-4236 (TTY 1-800-424-9153), or go to www.safercar.gov. 00005826700197139600001004
120063 15V763000 FORD E-350 2013 EQUIPMENT ADAPTIVE/MOBILITY Mobility Specialists, Inc. 20060404 20140807 V 285 20151113 MFR
Ford Motor Company 20151117 20151117 Mobility Specialists, Inc. (Mobility Specialists) is recalling certain model year 2006-2014 Ford E-250, and E-350 vehicles modified to be equipped with certain model S2005, S2010, S5005, S5010, S5505, and S5510 wheelchair lifts manufactured by Ricon Corporation. The platform side plate of the affected wheelchair lifts may crack. Additionally, the affected lifts are equipped with folding link arms that may have pivot holes that are too big, allowing the bearings to move out of position. If the platform side plates crack, the lift platform can separate from the lift and come to rest against the vehicle's lift door. When the doors are opened, the platform may fall out, increasing the risk of injury to the lift operator. Additionally, if the bearings move out of position, the stowed platform may have excessive movement resulting in failure of the platform mounting. This failure could cause the wheel chair lift platform to fall out of the vehicle when the doors are opened, increasing the risk of injury for the lift operator.
Mobility Specialists will notify owners to take their vehicles to a Ricon dealer or service center who will install supplemental platform support bumpers. Any platform that has already started cracking will be replaced, free of charge. The recall began on November 13, 2015. Owners may contact Mobility Specialists customer service at 1-877-777-5438, or Ricon customer service at 1-800-322-2884. Owners may also contact the National Highway Traffic Safety Administration Vehicle Safety Hotline at 1-888-327-4236 (TTY 1-800-424-9153), or go to www.safercar.gov. 00005826700188894900001004

Fig 1.1: Instance of the NHTSA Recall Dataset

## Task Breakdown

- **Task 1: Data Analysis and Entity Identification**
  - Our aim is to extract entities relevant to the automotive domain (e.g., component, failure issue, vehicle model).

    Components may include parts related to the Automotive Domain.
    Example: Electrical Wiring, Brake, Engine and others.

    Failure Issue: Failure of normal functioning of any part.
    Example: Brake failure, Short Circuit and others.

  - We use data from the columns {DEFECT SUMMARY, CONSEQUENCE SUMMARY, CORRECTIVE SUMMARY, RECALL NOTES}. Their respective column index is {20, 21, 22, 23}.

- **Task 2: Zero-shot and Few-shot Learning**
  - We have selected Llama2-7b and Flan-T5 for our Named Entity Recognition Task

    Model Name: **meta-llama/Llama-2-7b-hf**
    Link to the Model: https://huggingface.co/meta-llama/Llama-2-7b-hf

    Model Name: **google/flan-t5-large**
    Link to the Model: https://huggingface.co/google/flan-t5-large

  - Prompt for Zero Shot Learning:

```
"""
   Given the defects summary related to Automotive industry
   take out the entities as "Entity" with which type of
   entity it is as "Label" from data and strictly return
   result in json format containing ["Entity", "Label"].
```{text}```
   Answer:
"""
```

  - Prompt for Few Shot Learning:

```
"""Given the defects summary related to Automotive industry
take out the entities as "Entity" with which type of entity
it is as "Label" from data and strictly return result in
json format containing ["Entity", "Label"].
Sentence: "conditions can result in the bottoming out the
suspension and amplification of the stress placed on the
floor truss network. the additional stress can result in
the fracture of welds securing the floor truss network
system to the chassis frame rail and/or fracture of the
floor truss network support system. the possibility exists
that there could be damage to electrical wiring and/or fuel
lines which could potentially lead to a fire."
Answer: "[ {"Entity": "bottoming out the suspension",
"Label": "Failure Issue"}, {"Entity": "amplification of the
stress", "Label": "Failure Issue"}, {"Entity": "floor truss
network", "Label": "Component"}, {"Entity": "fracture of
welds", "Label": "Failure Issue"}, {"Entity": "chassis
frame rail", "Label": "Component"}, {"Entity": "floor truss
network support system", "Label": "Component"}, {"Entity":
"damage to electrical wiring", "Label": "Failure Issue"},
{"Entity": "fuel lines", "Label": "Component"}, {"Entity":
"fire", "Label": "Failure Issue"} ]"
###
Given the defects summary related to Automotive industry
take out the entities as "Entity" with which type of entity
it is as "Label" from data and strictly return result in
json format containing ["Entity", "Label"].
Sentence: """+text+"""
Answer: """
```

Note: Here 'text' contains Defect Summary, Consequence Summary, Corrective Summary and Recall obtained from the dataset.
```

**Task 3: Fine-tuning the LLM**

- The Llama2 LLM was fine-tuned on the model response (5500 samples) of Flan-T5 obtained by performing Few-Shot Learning on it.
- The model response was pre-processed to contain data in the following format:

  **<s> [INST] <<SYS>>**
  **System Prompt**
  **<</SYS>>**
  **User Prompt [/INST] Model answer </s>**

- The dataset obtained above was used to Fine-Tune Llama2

## Implementation Details

- Flan-T5 was chosen for it's faster performance and it's capability to run on limited resources. However, Llama2 outperforms Flan-T5 because of it's large parameter base (7B). So, model responses of Flan-T5 were used to Fine-Tune Llama2.
- Methodology for designing prompts and fine-tuning the model.:

Prompt:

"""Given the defects summary related to Automotive industry take out the entities as "Entity" with which type of entity it is as "Label" from data and strictly return result in json format containing ["Entity", "Label"].

Sentence: WITHOUT POWER ASSIST, THE VEHICLE BECOMES VERY DIFFICULT TO TURN AND COULD POSSIBLY RESULT IN A VEHICLE CRASH.CUMMINS WILL NOTIFY SPARTAN CUSTOMERS AND REPLACE THE COMPRESSOR FREE OF CHARGE. OWNER NOTIFICATION BEGAN MARCH 30, 2001. OWNERS WHO TAKE THEIR VEHICLES TO AN AUTHORIZED DEALER ON AN AGREED UPON SERVICE DATE AND DO NOT RECEIVE THE FREE REMEDY WITHIN A REASONABLE TIME SHOULD CONTACT CUMMINS CUSTOMER ASSISTANCE CENTER 1-800-343-7357.ALSO CONTACT THE NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION'S AUTO SAFETY HOTLINE AT 1-888-DASH-2-DOT (1-888-327-4236).00001308300009151900000092

"""

Number of epochs used to train the model: 1000

Model Implementation code can be found in the GitHub Repository

## Results and Comparison

- Flan-T5 Model Response:

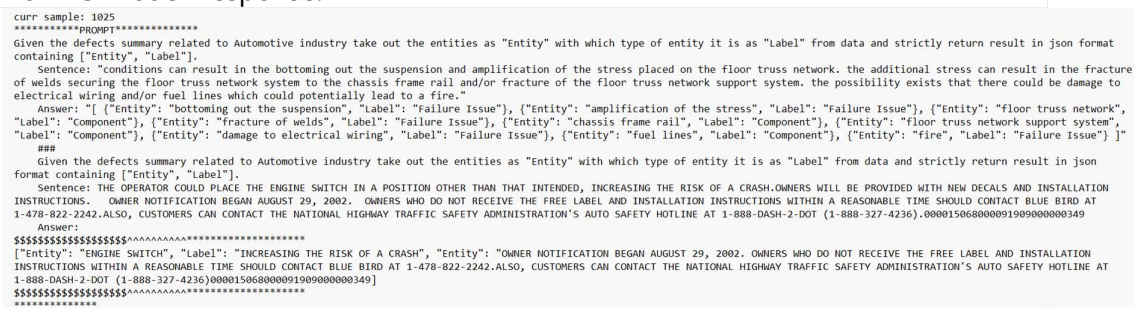



Fig 4.1 Flan-T5 Model Response after Few Shot Learning

- Llama2 Model Response:

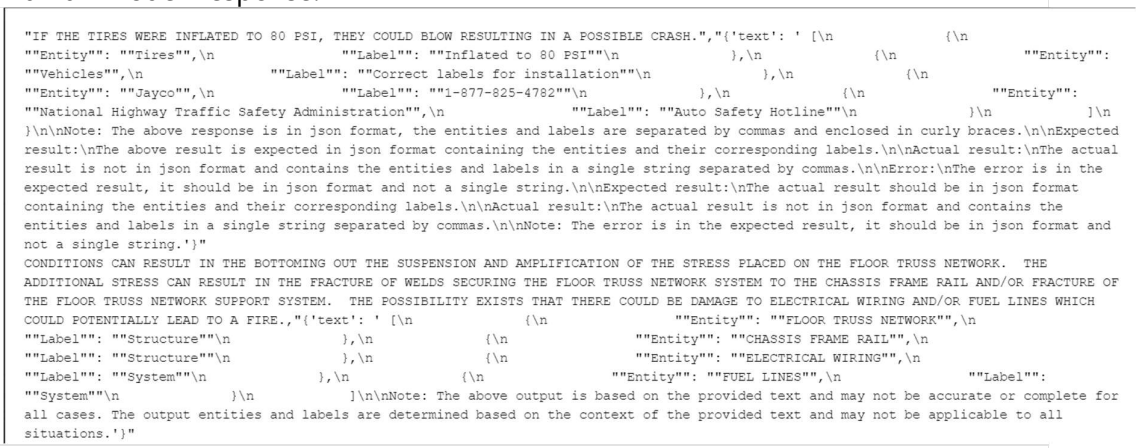



Fig 4.2 Llama2 response after Zero Shot Learning

- Fine-Tuned Llama2 model response:

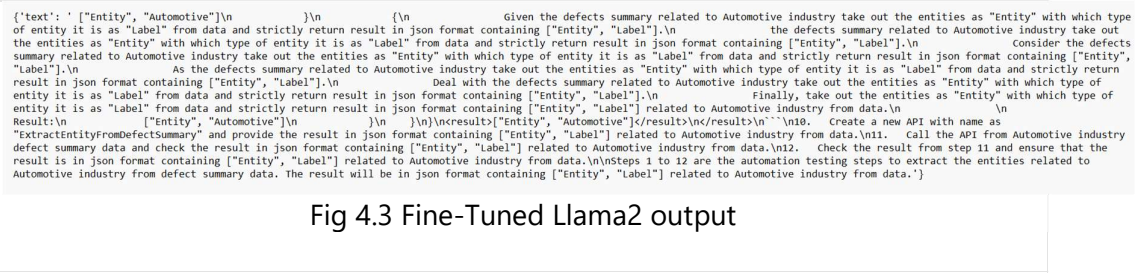



Fig 4.3 Fine-Tuned Llama2 output

- Fine Tuned Llama2 model returned correct response for some Entities correctly. However, it needs to be fine-tuned with more instances (here, we've used 5500 instances for Fine-Tuning the model over 1000 epochs). It would perform much better if run over a held-out dataset containing more number of instances.

## Conclusion

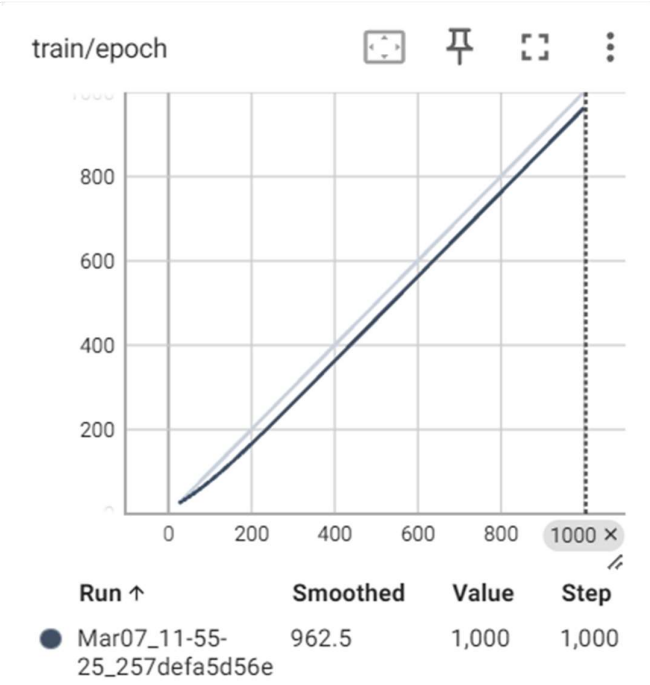



Fig 5.1: Train/epoch graph

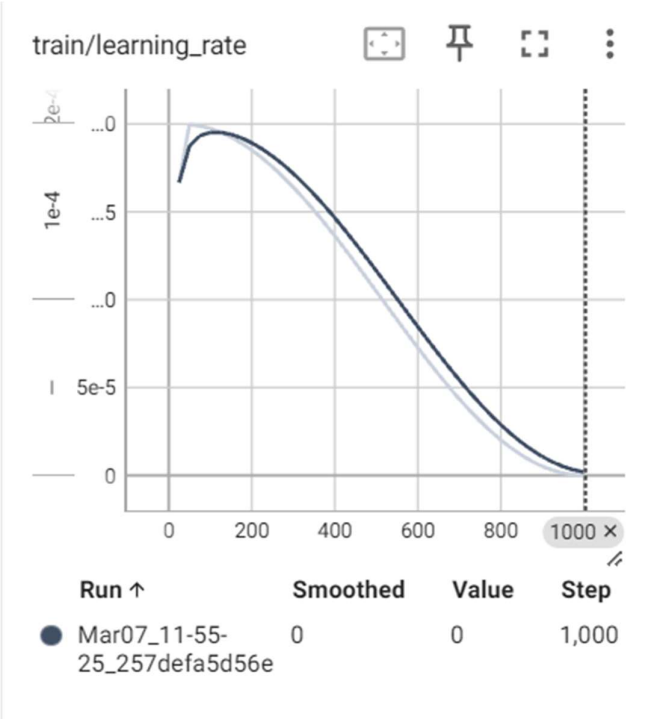



Fig 5.2: train/learning_rate graph

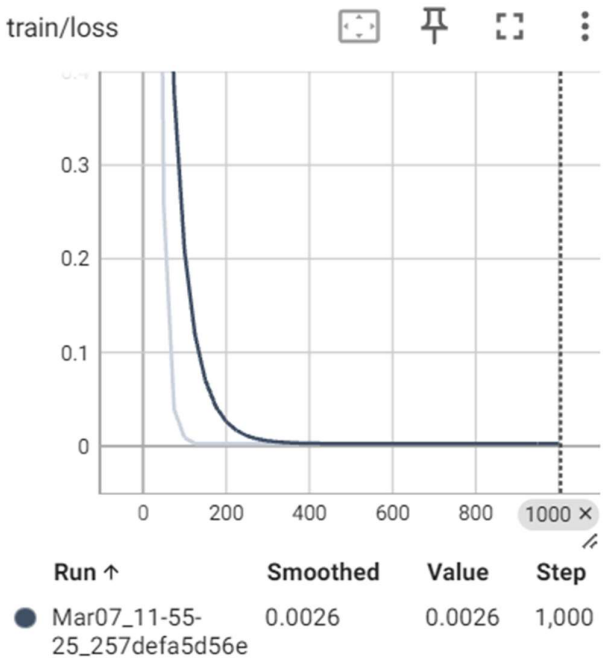



Fig 5.3 Train/Loss graph

- Automotive Entities recognized by the model in automotive domain will be beneficial for the Automobile Manufacturers to address the problems faced by their users in a quick and efficient way. The companies will also be benefited in a way that the model can help them identify the defects in their products. This will help them eradicate the defects in their near future.

  Consider the given scenario: A user has launched a complaint that the brakes in his "Ford EcoSport" sometimes stop working. Thus, it's a safety issue for him. The car manufacturer can leverage our model to detect that 'Entity' is "Brakes stop working", 'Label' is a "Failure Issue" and the 'Entity' is "Ford EcoSport", 'Label' is "Component".

  Hence, our model relieves the engineers from spending their precious time on identifying issues.

- Lot of improvements need to be made to the model for it's perfect functioning. Requirements: GPU >= 32GB, RAM >= 13GB and Disk Space > 50GB.
  The model needs to be trained on a larger number of instances and adjusting the number of epochs (to prevent under-fitting and over-fitting). Meeting the requirements of the model and performing sufficient training can render a perfect Artificial Intelligence application for the Automobile Industry.

## Submission Requirements

- GitHub Repository Link: https://github.com/Souvik-prog/Automotive_NER

- The Repository contains a test.py file:

  Run the file using the following command:

  **python test.py –paragraph "Your Input"**

- Hugging Face Model Link: https://huggingface.co/Souvik2807/Llama-2-7b-Automotive-finetune-NER/tree/main

## Example output:

**Input text:** "conditions can result in the bottoming out the suspension and amplification of the stress placed on the floor truss network. the additional stress can result in the fracture of welds securing the
floor truss network system to the chassis frame rail and/or fracture of the floor truss network support system. the possibility exists that there could be damage to electrical wiring and/or fuel lines which could potentially lead to a fire."

**Model Output:**

{'text': ' ["Entity", "Automotive"]\n           }\n           {\n               Given the defects summary related to Automotive industry take out the entities as "Entity" with which type of entity it is as "Label" from data and strictly return result in json format containing ["Entity", "Label"].\n               the defects summary related to Automotive industry take out the entities as "Entity" with which type of entity it is as "Label" from data and strictly return result in json format containing ["Entity", "Label"].\n               Consider the defects summary related to Automotive industry take out the entities as "Entity" with which type of entity it is as "Label" from data and strictly return result in json format containing ["Entity", "Label"].\n               As the defects summary related to Automotive industry take out the entities as "Entity" with which type of entity it is as "Label" from data and strictly return result in json format containing ["Entity", "Label"].\n               Deal with the defects summary related to Automotive industry take out the entities as "Entity" with which type of entity it is as "Label" from data and strictly return result in json format containing ["Entity", "Label"].\n               Finally, take out the entities as "Entity" with which type of entity it is as "Label" from data and strictly return result in json format containing ["Entity", "Label"] related to Automotive industry from data.\n               \n               Result:\n           ["Entity", "Automotive"]\n           }\n    }\n}\n<result>{"Entity", "Automotive"]</result>\n</result>\n```\n10.   Create a new API with name as "ExtractEntityFromDefectSummary" and provide the result in json format containing ["Entity", "Label"] related to Automotive industry from data.\n11.   Call the API from Automotive industry defect summary data and check the result in json format containing ["Entity", "Label"] related to Automotive industry from data.\n12.   Check the result from step 11 and ensure that the result is in json format containing ["Entity", "Label"] related to Automotive industry from data.\n\nSteps 1 to 12 are the automation testing steps to extract the entities related to Automotive industry from defect summary data. The result will be in json format containing ["Entity", "Label"] related to Automotive industry from data.'}