\documentclass{amsart}

\synctex=1

%=================================================================

%

\newcount\DraftStatus % 0 suppresses notes to selves in text

\DraftStatus=1 % TODO: set to 0 for final version

%=================================================================

%=================================================================

\usepackage{comment}

%=================================================================

%

\includecomment{JournalOnly}

\includecomment{ConferenceOnly}

\includecomment{TulipStyle}

%

%=================================================================

\input{preamble}

%=================================================================

%

\begin{document}

%

%=================================================================

%

\title[Taxi fare Prediction]{New York City Taxi fare Prediction}%

\author{Spoorthy Reddy JaruguA Short Running Title]{Title of This Paper}%

\author{Author 1}

\address[A.~1]{School of Computer Science,\\

Xi'an Shiyou University, Shaanxi 710065, China}%

\email[A.~1]{xxx@tulip.academy}

\author{Gang Li}

\address[A.~12]{Vellore Institute ofSchool of Information Technology - India

\\

Deakin University -, Geelong, VIC 3216, Australia}%

\email[A.~12]{jaruguspoorthyreddy@gmail.com}

gang.li@deakin.edu.au}

\author{Author 3}

\address[A.~3]{School of Information Technology \\

Deakin University \\

Vic 3125, Australia}%

\email[A.~3]{xxx@deakin.edu.au}

%\thanks{Thanks to \ldots}%

\subjclass{Artificial Intelligence}%

\date{2022-07-03\gitAuthorDate}%

\input{./tex/abstract}

\maketitle

\tableofcontents

\newpage

%=================================================================

\input{./tex/mainbody}

% ----------------------------------------------------------------

\newpage

\bibliography{tuliplab,yourbib}

% TODO: you should change this yourbib into a proper bib file name

\bibliographystyle{plainnat}

%=================================================================

\listoftodos

\end{document}

%=================================================================

\section{Introduction}\label{sec-intro}

Task is to predict the fare amount for a taxi ride in New York City. In table we have pickup and dropoff locations. We have to calculate the disatnce based on the date provided.

\begin{itemize}

\item The intersing \textcolor{orange}{characteristic} is how to calculate the

\textcolor{orange}{distance} from lattitude and longitude given.

\item Depedign upon the calcuated distance (trip duration), taxi fare is predicted

\end{itemize}

\section{Data Processing} \label{sec-preliminaries}

\begin{itemize}

\item

First step is to read the dataset from the CSV file

\item

Second print the both train and test Dataset

\item

Third check for NA values in the dataset

\bigskip

\smallskip

\vspace{0.75cm}

%\vspace{0.1cm}

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.55\textwidth, height=0.4\textheight]{NA-value-train.eps}

%\includegraphics[width=0.6\textwidth]{figures//example-basketball-projection.eps}\\

\caption{Listing missing vlaues in train dataset}\label{fig:Train}

\end{figure}

\newpage

\vspace{0.75cm}

%\vspace{0.1cm}

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.55\textwidth, height=0.35\textheight]{NA-value-test.eps}

%\includegraphics[width=0.6\textwidth]{figures//example-basketball-projection.eps}\\

\caption{Listing missing vlaues in test dataset}\label{fig:Test}

\end{figure}

\end{itemize}

Identified that NAN values are present in dropoff\_longitude and dropoff\_latitude

\begin{itemize}

\item

\smallskip

Removing the NAN values present in the train dataset by dropna command shown below and checking again for NAN values

\end{itemize}

\newpage

%\vspace{0.75cm}

%\vspace{0.1cm}

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics[width=0.5\textwidth,height=0.35\textwidth]{removing-nan.eps}\\

\caption{Removing NAN vlaues}\label{fig:nanvalue}

\end{figure}

\begin{itemize}

\item

Removing the data where pickup and dropoff locations are same (i.e pickup\_longitude and dropoff\_longitude; pickup\_latitude and dropoff\_latitude).

\item Checking for outliers by fixing the boundery of New York City

\begin{itemize}

\item minimum\_latitude is 40.573143,

\item minimum\_langitude is -74.252193,

\item maximum\_latitude is 41.709555,

\item maximum\_langitude is -72.986532

\end{itemize}

\item Removing outliers as they are identified

\end{itemize}

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics[width=0.6\textwidth,height=0.3\textwidth]{boundry.eps}\\

\caption{Removing NAN vlaues}\label{fig:Removing out of boundry data}

\end{figure}

As the data is very huge randomly we are selecting 10\% of data for the further process. The process of how data is ransomly selected is given below.

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics[width=0.6\textwidth,height=0.2\textwidth]{ten-percent-of-data.eps}\\

\caption{Selecting 10 percent of data}\label{fig:Selecting 10 percent of data}

\end{figure}

\begin{itemize}

\item

There are two paraters \textcolor{orange}{pickup - lattitude, longitude}, \textcolor{orange}{dropoff - lattitude, longitude}

\item Let us scatter plot the above parameters as pickup data and dropoff data

\newpage

\smallskip

\vspace{0.75cm}

%\vspace{0.1cm}

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.65\textwidth, height=0.4\textheight]{pickup.eps}

\caption{Pickup data}\label{fig:Pickup}

\end{figure}

\vspace{0.75cm}

%\vspace{0.1cm}

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.65\textwidth, height=0.4\textheight]{dropoff.eps}

\caption{Dropoff data}\label{fig:Dropoff}

\end{figure}

\end{itemize}

\begin{itemize}

\item

Print the count for passengers

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.3\textwidth, height=0.2\textheight]{passenger-count.eps}

\caption{Passenger count}\label{fig:Passenger count}

\end{figure}

\item

Print the maximum and minimum value in passenger and cleanign the data for maximum count of 6 passengers per ride

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.7\textwidth, height=0.3\textheight]{passenger-cleaning.eps}

\caption{Cleaned passenger data}\label{fig:Cleaned passangers data}

\end{figure}

\item

Print the count for passengers

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.65\textwidth, height=0.4\textheight]{passenger-visualization.eps}

\caption{Visualizing passenger data}\label{fig:Passenger count visualization}

\end{figure}

\end{itemize}

\newpage

\section{Data Extraction}

\begin{itemize}

\item To predict the taxi fare accurately we are extracting the

\begin{itemize}

\item Hour is calcuted to find weather its mid\_night\_trip or rush\_hour\_trip is noted

\item Day on which the passenger is picked upon

\item Month of trip

\item Year of travel

\end{itemize}

from the \textcolor{orange}{pickup\_datetime} coulmns

\item From the pickup\_month weather its snow\_season or not is noted

\item Finally trip\_diatance is calculated from pickup\_latitude, pickup\_longitude, dropoff\_latitude, dropoff\_longitude and stored it in trip\_distance

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.7\textwidth, height=0.25\textheight]{distance.eps}

\caption{Distance Calculation}\label{fig:Distance Calculation}

\end{figure}

\end{itemize}

\bigskip

\newpage

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.85\textwidth, height=0.5\textheight]{year-month-hour-week.eps}

\caption{Visualizing year, month, hour, weekday count}\label{fig:visualization year-month-hour-week}

\end{figure}

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.8\textwidth, height=0.5\textheight]{distance-fare.eps}

\caption{Visualizing trip\_distance and fare\_amount}\label{fig:visualization trip distance and fare}

\end{figure}

\newpage

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.9\textwidth, height=0.65\textheight]{heatmap.eps}

\caption{Visualizing heatmap}\label{fig:visualization heatmap}

\end{figure}

\newpage

\section{Model built and prediction} \label{sec-method}

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.8\textwidth, height=0.5\textheight]{linear-regression.eps}

\caption{Linear Regression}\label{fig:linear regression}

\end{figure}

\begin{itemize}

\item Bulit a Linear Regression model predict the fare\_amount of the trip in New York city

\item

\end{itemize}

\newpage

\section{Evaluating the model} \label{sec-experiment}

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.9\textwidth, height=0.55\textheight]{score.eps}

\caption{Evaluation Score}\label{fig:Evaluation score}

\end{figure}

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.9\textwidth, height=0.65\textheight]{predict-data.eps}

\caption{Visualizing predicted data}\label{fig:visualization predicted data}

\end{figure}

\begin{figure}[h!]

\centering

\selectcolormodel{rgb}

\includegraphics\*[width=0.9\textwidth, height=0.65\textheight]{testdata-prediction.eps}

\caption{Test data prediction}\label{fig:test data prediction}

\end{figure}

\newpage

\section{Conclusions} \label{sec-conclusions}

\begin{itemize}

\item

\smallskip

Fare prediction using latitude and longitude information is showcased.

\item

\smallskip

Additionally mid\_night\_trip, Rush\_hour\_trip, show\_season parameters are also considerd in fare calculation.

\item

\smallskip

The prediction model helps both passengers and drivers for effective fare prediction compared to conventional prediction

\end{itemize}

%\todo{Narrow down to a topic; Dig a hole; Fill the hole}

\todo{Formula for Introduction}

%\gangli{``narrow in on topic'' reminds you

%that readers and reviewers only know that this is a AI or HTM research paper (and maybe have read the title/abstract).

%You need to help them figure out what topic and area of research paper this is.

%You \_don't\_ need to wax poetic about the topic's importance.}

%\gangli{`dig a hole'' reminds you that

%you need to convince the reader that there's a problem with the state of the world.

%Prior work may exist but it's either missing something important or there's a missing opportunity.

%The reader should be drooling for a bright future just out of reach.}

%\gangli{``fill the hole'' reminds you to show the reader

%how and why the paper they're reading will fix these problems and deliver us into a better place.

%You don't need a whirlwind summary of the technical details,

%but you need readers convinced (and in a good mood) to keep reading.}

\gangli{A good paper introduction is fairly formulaic.

If you follow a simple set of rules,

you can write a very good introduction.

The following outline can be varied.

For example,

you can use two paragraphs instead of one,

or you can place more emphasis on one aspect of the intro than another.

But in all cases,

all of the points below need to be covered in an introduction,

and in most papers,

you don't need to cover anything more in an introduction.}

%\todo{The importance of the area}

%\blindtext

\todo{Motivation}

At a high level,

what is the problem area you are working in and why is it important?

It is important to set the larger context here.

Why is the problem of interest and importance to the larger community?

%\todo{The problems faced by most current methods}

%\blindtext

\todo{What is the specific problem considered in this paper?}

This paragraph narrows down the topic area of the paper.

In the first paragraph you have established general context and importance.

Here you establish specific context and background.

%\todo{What can be addressed by existing methods; Why those problems are challenges to existing methods?}

%\blindtext

\todo{Contribution}

"In this paper, we show that ...".

This is the key paragraph in the intro - you summarize,

in one paragraph,

what are the main contributions of your paper given the context

you have established in paragraphs 1 and 2.

What is the general approach taken?

Why are the specific results significant?

This paragraph must be really good.

You should think about how to structure these one or

two paragraph summaries of what your paper is all about.

If there are two or three main results,

then you might consider itemizing them with bullets or in test.

\begin{itemize}

\item e.g., First ...

\item e.g., Second ...

\item e.g., Third ...

\end{itemize}

If the results fall broadly into two categories,

you can bring out that distinction here.

For example, "Our results are both theoretical and applied in nature.

(two sentences follow, one each on theory and application)"

%\todo{What provides the motivation of this work? What are the research issues? What is the rationale of this work? }

%\blindtext

\todo{At a high level what are the differences in what you are doing, and what others have done? }

Keep this at a high level,

you can refer to a future section where specific details and differences will be given.

But it is important for the reader to know at a high level,

what is new about this work compared to other work in the area.

%\todo{What we have done and what are the contributions.}

%\blindtext

\todo{A roadmap for the rest of the paper}

"The remainder of this paper is structured as follows..."

Give the reader a roadmap for the rest of the paper.

Avoid redundant phrasing,

"In Section 2, In section 3, ... In Section 4, ... " etc.

\gangli{A few general tips:

Don't spend a lot of time into the introduction

telling the reader about what you don't do in the paper.

Be clear about what you do do.

Does each paragraph have a theme sentence that sets the stage for the entire paragraph? Are the sentences and topics in the paragraph all related to each other?}

\gangli{Does each paragraph have a theme

sentence that sets the stage for the entire paragraph?

Are the sentences and topics in the paragraph all related to each other?}

\gangli{Do all of your tenses match up in a paragraph?}

Test citation~\cite{BL12J01}.

\begin{JournalOnly}

and~\citep{BJL11J01} or~\citet{BJL11J01}.

\end{JournalOnly}

This is for~\cref{tbl:overall-experiments},

\todo[fancyline]{Testing.}

and this is for~\cref{sec-conclusions}.

\todo[noline]{A note with no line back to the text.}%

\gangli{This is comment from Gang.}

\qwu{Response from QW}

Number:

\num{123}.

\numlist{10;30;50;70},

\numrange{10}{30},

\SIlist{10;30;45}{\metre},

and

\SI{10}{\percent}

\missingfigure[figcolor=white]{Testing figcolor}

\begin{ConferenceOnly}

We have \SI{10}{\hertz},

\si{\kilogram\metre\per\second},

the range: \SIrange{10}{100}{\hertz}.

$\nicefrac[]{1}{2}$.

\missingfigure{Make a sketch of the structure of a trebuchet.}

\end{ConferenceOnly}

For~\cref{eq:test},

as shown below:

\begin{equation}\label{eq:test}

a = b \times \sqrt{ab}

\end{equation}

\blindmathpaper

\section{Preliminaries} \label{sec-preliminaries}

\blindtext

\gliMarker %TODO: GLi Here

\section{Method} \label{sec-method}

\blindtext

\blindlist{itemize}[3]

\blinditemize

\blindenumerate

\blindmathtrue

\blindmathfalse

\blinddescription

\qwuMarker %TODO: QWu Here

\section{Experiment and Analysis} \label{sec-experiment}

\begin{table} \centering

\caption{Precision Comparison on Event Detection Methods}

\label{tbl:overall-experiments}

\begin{tabular}{cccc}

\toprule

% after \\: \hline or \cline{col1-col2} \cline{col3-col4} ...

& OR Event Detection & AC Event Detection & TC Event Detection \\

\midrule

precision & 0.83 & 0.69 & 0.46 \\

recall & 0.68 & 0.48 & 0.36 \\

F-score & 0.747 & 0.57 & 0.4 \\

\bottomrule

\end{tabular}

\end{table}

\section{Conclusions} \label{sec-conclusions}

\blindtext

\section\*{Acknowledgement}

\lipsum[1]

The authors would like to thank \ldots