Avoiding Stress Driving: Online Trip Recommendation from Driving Behavior Prediction

Authors: Rohit Verma, Bivas Mitra and Sandip Chakraborty

Indian Institute of Technology Kharagpur
Past Few Years
Increasing Number of Road Accidents  →  Worried Cab Companies
Past Few Years
Increasing Number of Road Accidents \Rightarrow Worried Cab Companies

Several Studies(1,2) \Rightarrow Driving Stress

Increasing Number of Road Accidents → Worried Cab Companies

Several Studies (1,2) → Driving Stress

Dangerous Driving Behavior

Consider this Scenario

8th trip
Consider this Scenario

8th trip
Consider this Scenario

8th trip
Consider this Scenario

8th trip
Consider this Scenario

8th trip

Dangerous Driving Behavior
Objective

• Develop a recommendation system, which based on the driver’s personality traits, predicts if a new trip will lead to stress resulting in dangerous driving behavior and recommends if the driver should Accept/Reject the trip.
Breaking down the Objective

• **Stress Model**: Develop a model to compute stress taking into account a driver’s personality traits.

• **Driving Behavior Model**: Quantify dangerous driving behavior.

• **Behavior Prediction Model**: Develop a model to predict driving behavior from stress.

• **Recommender**: Recommend the driver to Accept or Reject a trip.
How to detect driver stress?

Surveys


How to detect driver stress?

**Surveys**


**Physiological Sensors**


How to detect driver stress?

Surveys


Physiological Sensors


How to detect driver stress?


Can we use driving data to predict driver stress?
Stress Model

- We use a Neural Network to train over the drivers.
- For each driver we classify the stress into 3 levels (No, Medium and High).
Stress Model

- We use a Neural Network to train over the drivers.
- For each driver we classify the stress into 3 levels (No, Medium and High).
- Each driver has different set of **personality traits** which should be addressed by the model.
## Features Used

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of trips ($n_T$)</td>
<td>Number of trips the driver has covered including the current one.</td>
</tr>
<tr>
<td>Trip time covered ($t_T$)</td>
<td>Time for which the driver was driving starting from the first trip.</td>
</tr>
<tr>
<td>Trip distance covered ($d_T$)</td>
<td>Distance for which the driver was driving starting from the first trip.</td>
</tr>
<tr>
<td>Rest time ($r_T$)</td>
<td>Time for which the driver had taken rest after the last trip.</td>
</tr>
<tr>
<td>Time of the day ($z$)</td>
<td>Divided into 4 time zones. 6 AM -10 AM(0), 10 AM- 4 PM(1), 4 PM - 10 PM(2), 10 PM - 6 AM(3)</td>
</tr>
<tr>
<td>Congestion ($C$)</td>
<td>Calculated from the trajectory data using existing models</td>
</tr>
<tr>
<td>Road Type ($r$)</td>
<td>City (0), Highway (1), Rural (2). If multiple road types are on the same trip, then the score is calculated as the weighted average over the distance for which each type of road was driven on.</td>
</tr>
</tbody>
</table>
Stress Model: Multi-task Learning

- To address personalization on stress we utilize Multi-Task Learning.
- Task – Driver
- Class – Stress Label
- The objective of the model is to conduct a robust learning by
  - Shared learning: learning features of one driver (one task) using the related features of other drivers (related tasks)
  - Task-specific learning: the model is specialized to learn the characteristics leading to the stress label of the specific driver
Training and Evaluation

• Dataset Used: HCILab Dataset[1]
  • Drivers: 10 (3 female and 7 male)
  • Sensors: IMU sensors, GPS, ECG, SCR, Temperature, Heart Rate(HR) and HR Variability

Training and Evaluation

• Dataset Used: HCILab Dataset[1]
  • Drivers: 10 (3 female and 7 male)
  • Sensors: IMU sensors, GPS, ECG, SCR, Temperature, Heart Rate(HR) and HR Variability

• Ground Truth Generation:
  • We use the technique given by Keshan et. al.[2] to compute stress from the physiological sensor data.

Training and Evaluation

• Dataset Used: HCILab Dataset[1]
  • Drivers: 10 (3 female and 7 male)
  • Sensors: IMU sensors, GPS, ECG, SCR, Temperature, Heart Rate(HR) and HR Variability

• Ground Truth Generation:
  • We use the technique given by Keshan et. al.[2] to compute stress from the physiological sensor data.

• Evaluation:
  • We divide the data into 60-20-20% for training, validation and testing
  • We also implement a Single Task Learning model to train each driver in isolation

Results

Comparison with STL

- The MTL-NN approach has an **AUC of 0.931 compared to 0.794** of the STL approach.
Results

Comparison with STL

• The MTL-NN approach has an **AUC of 0.931 compared to 0.794** of the STL approach.

![AUC Comparison Graph](.graph.png)

Comparison with Existing models

• We compare with three models,
  - **Salai et. al.** – Developed an algorithm to detect stress from HRV
  - **Shi et. al.** – Employed SVM to detect stress using multiple physiological sensors
  - **Singh et. al.** – Developed a NN Model to compute stress using multiple physiological sensors

<table>
<thead>
<tr>
<th>Model</th>
<th>Area Under Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salai et. al. [30]</td>
<td>0.75</td>
</tr>
<tr>
<td>Shi et. al. [31]</td>
<td>0.62</td>
</tr>
<tr>
<td>Singh et. al. [32]</td>
<td>0.94</td>
</tr>
<tr>
<td>Our Model</td>
<td>0.93</td>
</tr>
</tbody>
</table>
Driving Behavior Score: Speed Profile

• Statutory speed limits ($\mathcal{L}$) are defined for all countries with some tolerance ($\mathcal{T}$).

• We give the score as:

$$
\mathcal{V} = \begin{cases} 
0, & s < \mathcal{L} \ (\text{safe}) \\
\frac{s - \mathcal{L}}{\mathcal{T}}, & \mathcal{L} \leq s < \mathcal{L} + \mathcal{T} \ (\text{moderate}) \\
1, & s \geq \mathcal{L} + \mathcal{T} \ (\text{dangerous})
\end{cases}
$$

where $s$ is the average speed of the car.
Driving Behavior Score: Interaction with PoCs

- We observe driving behavior while interacting with speed breaker or potholes

- Jerk observed during interaction is usually a measure of discomfort during such interaction which is calculated as

$$J = \frac{da(t)}{dt}$$

where $a(t)$ is the acceleration at time $t$

- We say the interaction is dangerous if the jerk is critical and give the score as;

$$J = \begin{cases} 
0, & J > -9.9 \text{m/s}^3 \\
1, & J < -9.9 \text{m/s}^3
\end{cases}$$

The value of $-9.9 \text{ m/s}^3$ for critical jerk was given by Nygard et. al[1].

Driving Behavior Score: Dangerous Maneuvers

- We extract six types of dangerous maneuvers
  - Weaving
  - Swerving
  - Side-Slipping
  - Fast U-turn
  - Sharp turn
  - Sudden brake

- We use a technique given by Yu et. al.[1];
  - Utilize inertial sensor data and SVM to detect the dangerous maneuvers
  - The model provides a tuple $\mathcal{M}$ of size six
  - If a dangerous maneuver is detected that index is set as 1 otherwise 0

Driving Behavior Score: Overall Score

- We have three different scores
  - The speed profile score $\mathcal{V}$
  - The interaction score $\mathcal{I}$
  - The dangerous maneuver tuple $\mathcal{M}$
Driving Behavior Score: Overall Score

• We have three different scores
  • The speed profile score $\mathcal{V}$
  • The interaction score $\mathcal{I}$
  • The dangerous maneuver tuple $\mathcal{M}$

• The overall score is given as:

$$
\mathcal{D} = \frac{1}{8} \sum \left( \mathcal{V} + \mathcal{I} + \sum_{i=1}^{6} \mathcal{M}_i \right)
$$
Driving Behavior Score: Overall Score

• We have three different scores
  • The speed profile score $V$
  • The interaction score $I$
  • The dangerous maneuver tuple $M$

• The overall score is given as:

$$D = \frac{1}{8} \sum \left( V + I + \sum_{i=1}^{6} M_i \right)$$

Following this we compute the stress and driving behavior score for two different datasets and try to observe if they have any relation.
Datasets

• UAH Driverset Data
  • 6 drivers
  • Inertial sensors, GPS and video data
  • 500 minutes of total driving data

• In-house Dataset
  • 8 drivers
  • Inertial sensors, GPS and video data
  • Data of 1700 trips for a duration of 5 months
In-house Dataset: Data Collection System

Data Collection Device

<table>
<thead>
<tr>
<th>Driver</th>
<th>Age Range</th>
<th>Vehicle Model</th>
<th>Fuel Type</th>
<th>Number of Trips</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>20-30</td>
<td>Tata Indigo</td>
<td>Diesel</td>
<td>282</td>
</tr>
<tr>
<td>D2</td>
<td>20-30</td>
<td>Hindustan Ambassador</td>
<td>Diesel</td>
<td>321</td>
</tr>
<tr>
<td>D3</td>
<td>20-30</td>
<td>Hindustan Ambassador</td>
<td>Diesel</td>
<td>325</td>
</tr>
<tr>
<td>D4</td>
<td>30-40</td>
<td>Tata Sumo</td>
<td>Diesel</td>
<td>72</td>
</tr>
<tr>
<td>D5</td>
<td>30-40</td>
<td>Maruti Suzuki Dzire</td>
<td>Diesel</td>
<td>67</td>
</tr>
<tr>
<td>D6</td>
<td>20-30</td>
<td>Tata Indigo</td>
<td>Diesel</td>
<td>80</td>
</tr>
<tr>
<td>D7</td>
<td>30-40</td>
<td>Maruti Suzuki Dzire</td>
<td>Diesel</td>
<td>75</td>
</tr>
<tr>
<td>D8</td>
<td>20-30</td>
<td>Maruti Suzuki Alto</td>
<td>Petrol</td>
<td>405</td>
</tr>
</tbody>
</table>
Correlating Stress and Driving Behavior Score

Driving behavior score with respect to the stress value for all the drivers

Kendall’s tau coefficient value for correlation between stress and driving behavior
(We obtain a mean correlation of 0.83 and p value of $2.99 \times 10^{-10}$)

It is evident that increase in stress deteriorates the driving behavior implying strong correlation between the two which is strengthened by the Kendall’s tau coefficient.
Correlating Stress and Driving Behavior Score

However this correlation can be spurious!

Driving behavior score with respect to the stress value for all the drivers

Kendall’s tau coefficient value for correlation between stress and driving behavior
(We obtain a mean correlation of 0.83 and p value of $2.99 \times 10^{-10}$)

It is evident that increase in stress deteriorates the driving behavior implying strong correlation between the two which is strengthened by the Kendall’s tau coefficient.
Causality Analysis

• We perform causality analysis to ensure the correlation is not spurious.

• Stress (S) $\rightarrow$ Treatment

• Driving Behavior Score ($D$) $\rightarrow$ Response

• We introduce confounding variables ($Z$) which might impact $D$ instead of $S$.

• The impact is given as Average Treatment Effect (ATE);

$$ATE = \mathbb{E} \left( \frac{D(u) - D(v)}{X(u) - X(v)} \right)_{(u,v) \in P}$$

where $X = S \cup Z$,

$u$ and $v$ are similar scenarios only differing w.r.t one of the treatment or confounding variables.
Causality Analysis

- We identify four confounding variables
  - Weather conditions (W)
  - Previous driving score (P)
  - Day of the week (Q)
  - Special Occasion (O)

<table>
<thead>
<tr>
<th></th>
<th>W</th>
<th>P</th>
<th>Q</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau$</td>
<td>0.68</td>
<td>0.56</td>
<td>0.18</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Kendall’s tau coefficient for the Confounding Variables w.r.t. $D$
Causality Analysis

- We identify four confounding variables
  - Weather conditions (W)
  - Previous driving score (P)
  - Day of the week (Q)
  - Special Occasion (O)

- We compute the ATE over $\mathcal{D}$ w.r.t. $S$, $W$ and $P$

<table>
<thead>
<tr>
<th>Variable</th>
<th>$S$</th>
<th>$W$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATE</td>
<td>0.62</td>
<td>0.23</td>
<td>0.21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$W$</th>
<th>$P$</th>
<th>$Q$</th>
<th>$O$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau$</td>
<td>0.68</td>
<td>0.56</td>
<td>0.18</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Kendall’s tau coefficient for the Confounding Variables w.r.t. $\mathcal{D}$
Causality Analysis

• We identify four confounding variables
  • Weather conditions (W)
  • Previous driving score (P)
  • Day of the week (Q)
  • Special Occasion (O)

• We compute the ATE over $\mathcal{D}$ w.r.t. $S$, $W$ and $P$

<table>
<thead>
<tr>
<th>Variable</th>
<th>$S$</th>
<th>$W$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATE</td>
<td>0.62</td>
<td>0.23</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Kendall's tau coefficient for the Confounding Variables w.r.t. $\mathcal{D}$

<table>
<thead>
<tr>
<th>Variable</th>
<th>$W$</th>
<th>$P$</th>
<th>$Q$</th>
<th>$O$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau$</td>
<td>0.68</td>
<td>0.56</td>
<td>0.18</td>
<td>0.23</td>
</tr>
</tbody>
</table>

A high Average Treatment Effect for stress ensures the non-spurious correlation
Driving Score Prediction

• Deriving from the high correlation results, we develop a prediction model for driving behavior from driving stress.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSE</td>
</tr>
<tr>
<td>Simple Linear Regression</td>
<td>0.0103</td>
</tr>
<tr>
<td>Linear SVR</td>
<td>0.0128</td>
</tr>
<tr>
<td>SVR (RBF kernel)</td>
<td>0.0131</td>
</tr>
<tr>
<td>Decision Tree Regression</td>
<td>0.0110</td>
</tr>
</tbody>
</table>

• A simple linear regression model is employed for prediction which gives the best result as shown in the table.
Trip Recommendation

New trip arrives
• System computes the required parameters

Compute Driver Stress
• Use the historical and current trip information in the Stress Model

Estimate Driving Score
• Use the prediction model to estimate the driving score if the next trip is accepted.

Compare with a pre-set threshold
• If the Driving score is above a threshold, ask the driver not to take the trip
Evaluation

• We evaluate the system over 7 drivers in the in-house dataset for a week.

• We start recommending after the third trip

• The threshold for Driving Score was set as 0.6
Results: Impact of Recommendation

- All drivers observe gain accepting recommendation.
- Driving score increases (deteriorates) when driver starts rejecting multiple recommendations.
Result: Impact on income

- Change is less than 25% for all the drivers. Moreover, most being increased which is a gain.

- 20% change would seem a major setback for some, but putting safety as the primary concern this is expected.
Conclusion

• We have used driving data to develop a personalized model to compute driver’s stress.

• We establish a strong quantitative relationship between driving behavior and driving stress.

• We provide a model to predict driving behavior from stress.

• We utilize these to develop a trip recommendation system for drivers.

• We could also use the system for
  • Full day roster generation
  • Award system for better drivers
  • Driver recruitment based on how they cope in different scenarios

• We still need to look into how some non-quantifiable confounding variables like car condition, family issues, etc. can be utilized by the system.
Acknowledgements

• Google – For supporting my travel.
Thank you!

Follow the work of Complex Network Research Group (CNeRG), IIT KGP at:
Web: http://www.cnerg.org
Facebook: https://web.facebook.com/iitkgpcnerg
Twitter: https://www.twitter.com/cnerg