



Natural Language Processing: subreddit 'cars' vs 'teslamotors'

By Alex Lau
January 2020



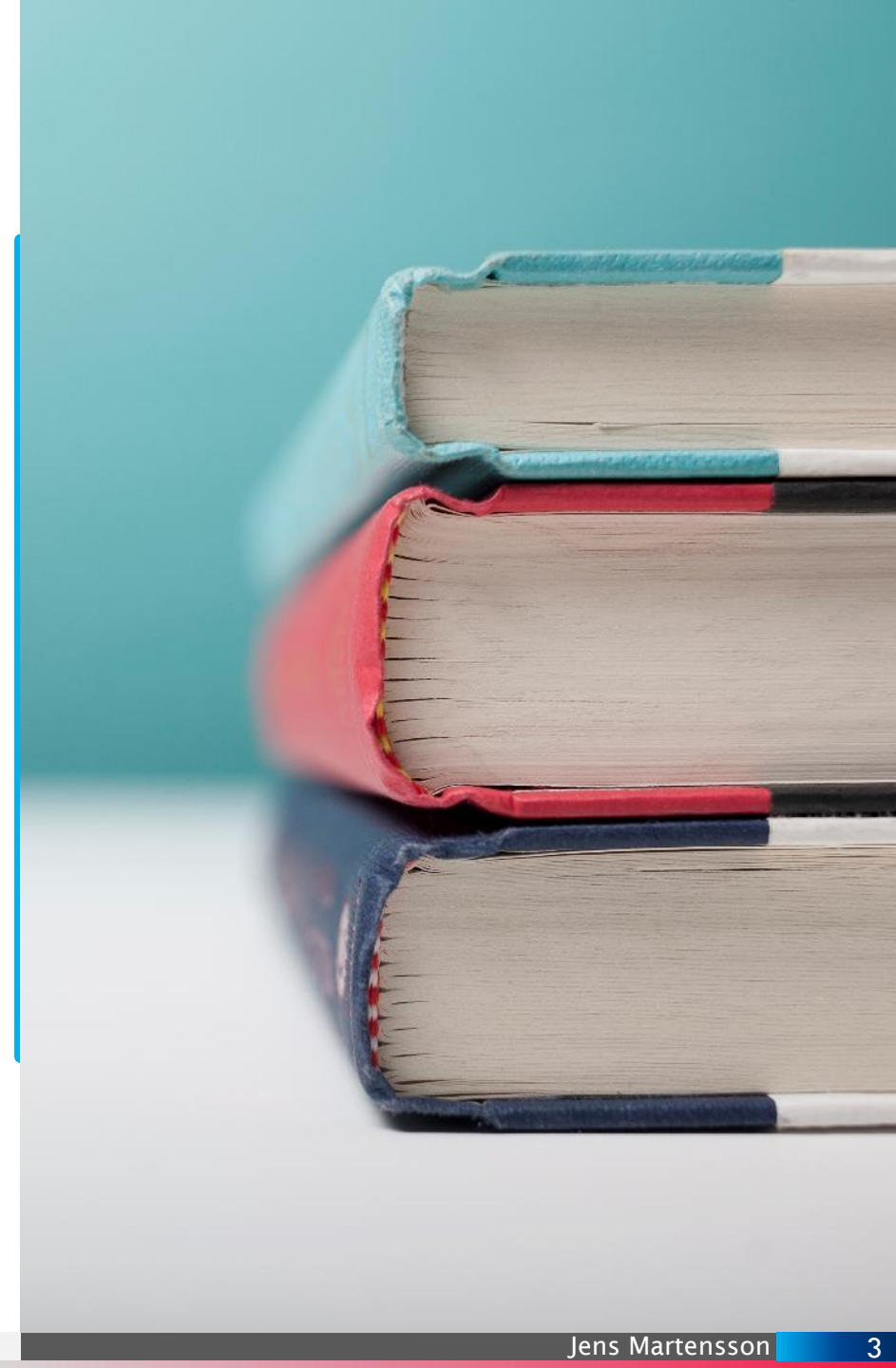
It is time for the
world to transition
to sustainable
energy

Tesla Motors is providing
solutions



Transportation Status Quo is Disastrous!

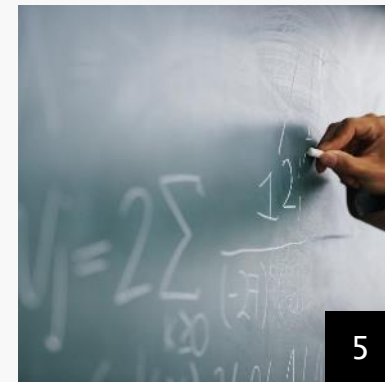
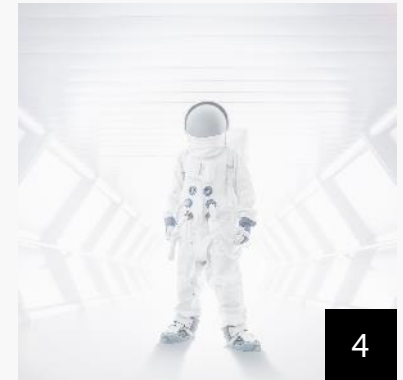
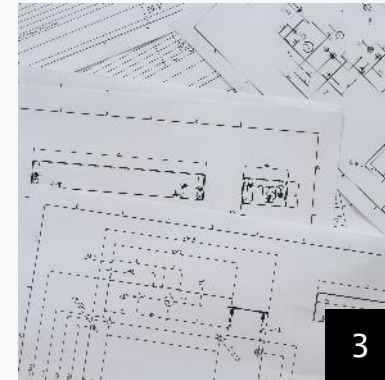
- Motor vehicles collectively cause 75 percent of carbon monoxide pollution in the U.S.
- Transportation causes 27 percent of greenhouse gas emissions
- The U.S. has 30 percent of the world's automobiles, yet it contributes about half of the world's emissions from cars
- On-road vehicles cause one-third of the air pollution that produces smog in the U.S.



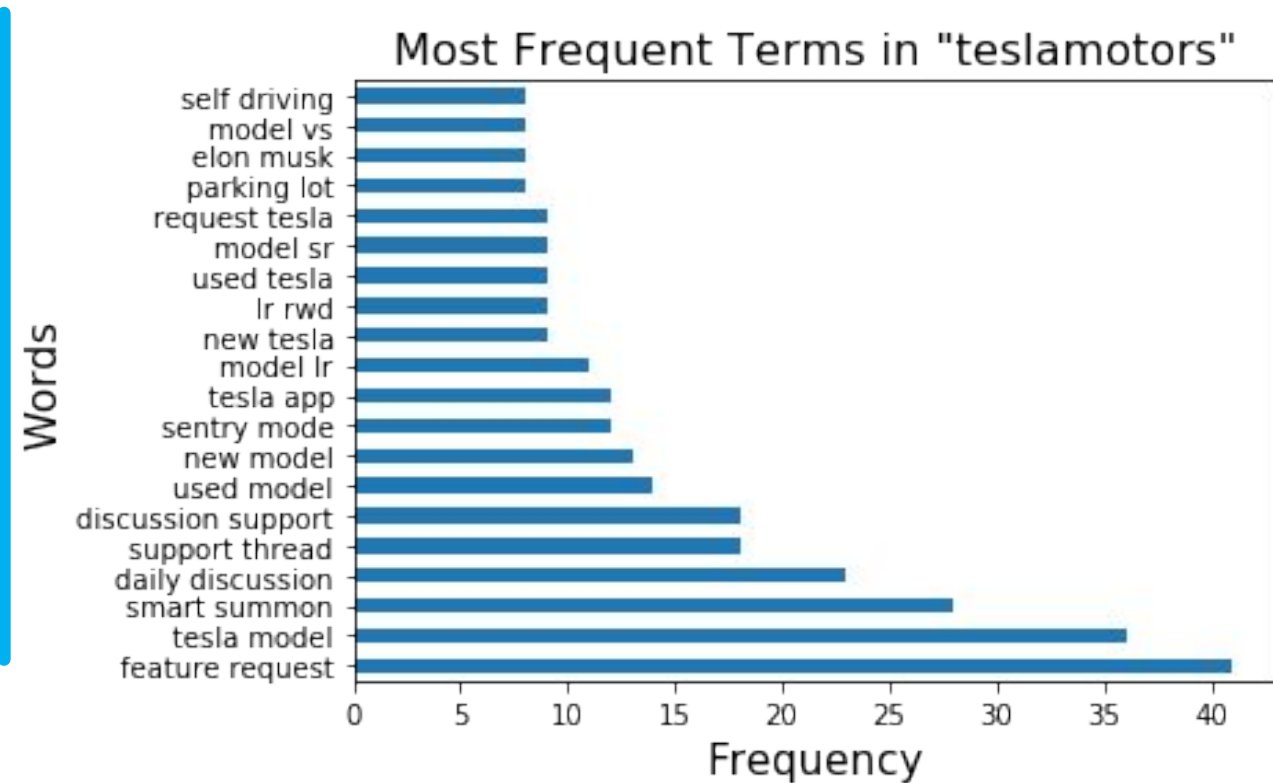
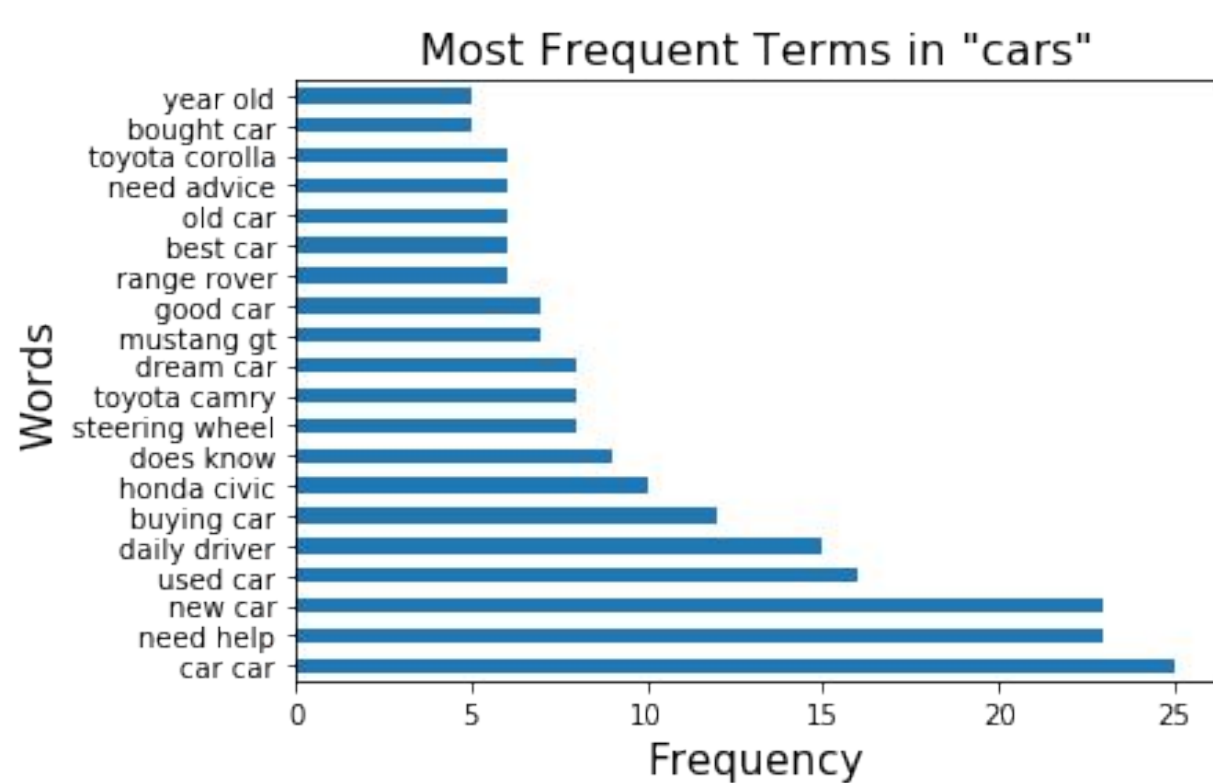
Problem Statement

What key words in titles differentiate users who post on subreddit 'cars' versus 'teslamotors'? How can we leverage this information to influence and educate the population to transition to sustainable energy?

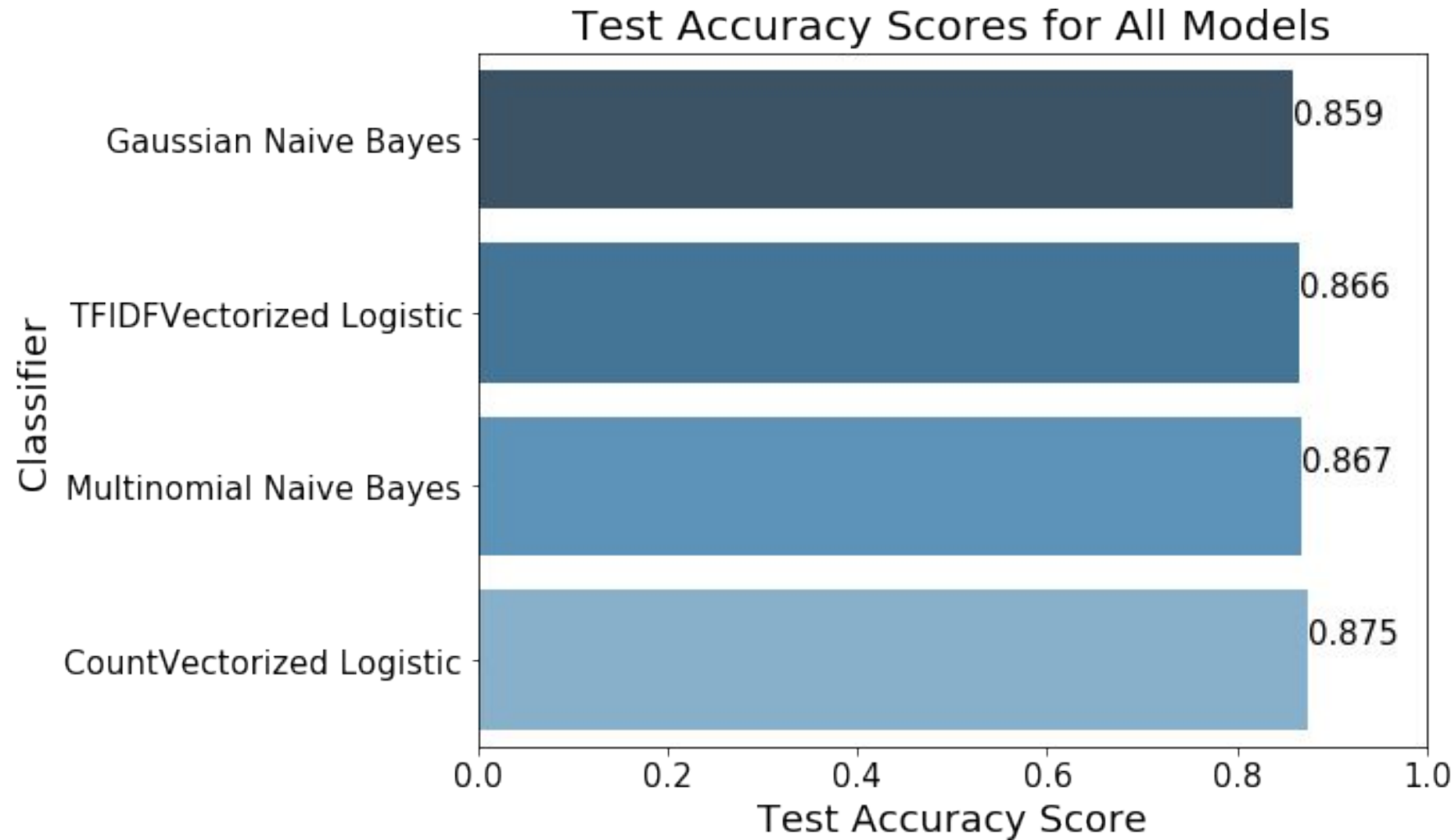
- Goal: Accurately distinguish titles between subreddit 'cars' and 'teslamotors', and outperform the baseline of 55.45%
- Method:
 - Count Vectorized Logistic Regression
 - TF-IDF Vectorized Logistic Regression
 - Multinomial Naïve Bayes Classifier
 - Gaussian Naïve Bayes Classifier
- Success Metrics: Accuracy Scores



EDA – Classification Bigram Comparisons



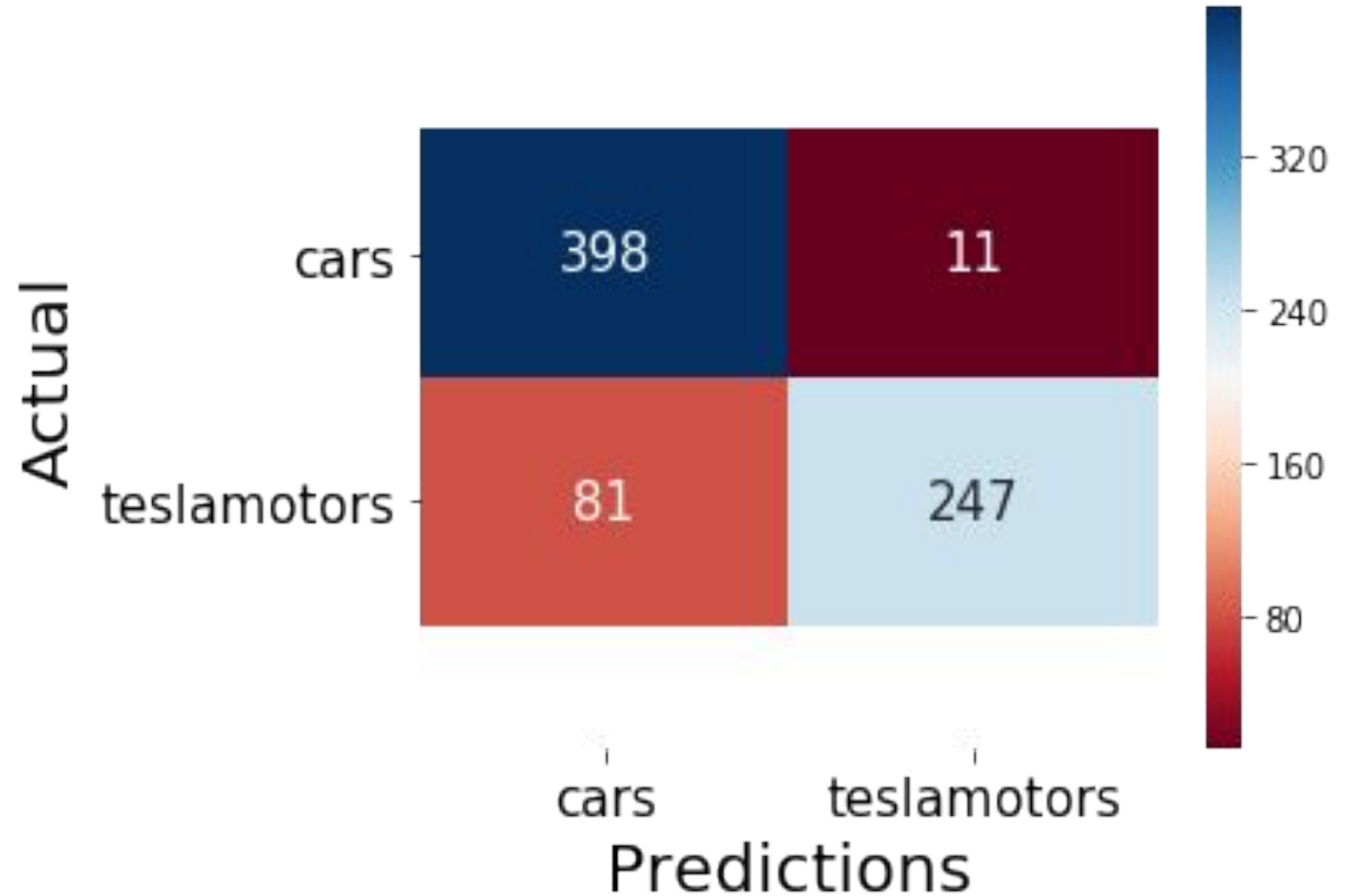
Performance – Count Vectorized Logistic Regression



Other Classification Metrics

- Accuracy: 87.52%
- Recall: 75.30%
- Specificity: 97.31%
- Precision: 95.74%

Confusion Matrix



Top 10 Words for 'teslamotors' Excluding Make and Model

Features	Coefficients	Log-Odds
autopilot	2.17	8.79
v10	1.90	6.69
summon	1.82	6.17
mode	1.65	5.19
elon	1.63	5.11
deliveri	1.59	4.93
app	1.42	4.16
fsd	1.41	4.11
updat	1.37	3.93
charg	1.35	3.87

Conclusion

Count-Vectorized Logistic Regression

- Accuracy score of 87.52% outperforms our baseline of 55.45%
- We have identified key words in titles of 'teslamotors' that closely relate to the innovations lead by Tesla
 - **Autopilot** – A driver assistance safety feature that utilizes cameras, automatic emergency breaking
 - **V10** – Tesla's biggest software version update that includes
 - Entertainment interface with Netflix, Hulu, Disney+
 - Smart Summon that allows the car to drive to you in a parking lot
 - Spotify interface for individual account access
 - **Summon** – The ability to move your car forward and backward with your phone like a remote control
- We recommend raising awareness of the life-saving technologies that exist and encourage people to break from the status quo of gas powered vehicles



Resources

Tesla V10 Software:

<https://www.tesla.com/blog/introducing-software-version-10-0>

Car Pollution Statistics:

<https://auto.howstuffworks.com/air-pollution-from-cars.htm>

✉ Alex.lau14@outlook.com

