SOLUTION PRESENTATION

* Document what you have done
* Create a nice presentation
  + Make sure you highlight the big picture first
* Explain why your solution achieves the business objective
* Don’t forget to present interesting points you noticed along the way
  + Describe what worked and what did not
  + List your assumptions and your system’s limitations

CONTEXT

* Present understanding of the business problem
* Frame problem such that it is understandable for different stakeholder
* Should cover brief resuime of 5.1: Framing the problem
* Give statistics about the current situation
* Present objectives and benefits

HYPOTHESIS

* Document/communicate the hypotheses in terms of:
  + What can be predicted/optimized
  + What are actions that can be taken?
  + What are the possible outcomes?
* Communicate your hypothesis testing strategy to the wide audience
  + Null hypothesis
  + Outcomes and types of errors
  + P-values and confidence intervals
  + Multiple testing and p-hacking

DATA

* Document and present leveraged data sources used to create the dataset
* Cover a brief resume of Data Acquisition, Data Exploration and Data Preparation
* Profile and present the data before and after going through acquisition, exploration and preparation

MODEL

* Document and present the type of model
  + Supervised, unsupervised, semi-supervised, reinforcement
  + Classification (binary, multi-class, multi-label, multi-output), regression
* Cover brief resume of 5.5 Modelling
* Modelling approach
  + Algorithms, benchmarks, fine-tuning, phases
* Model Evaluation
  + Training, validation, and testing dataset splitting strategy
  + Number of instances and percentages
  + Utilized evaluation and thought process

RESULTS

* Document and communicate evaluation metric results of each phase in modelling approach
* Communicate performance of model
* Compare the result with SOTA and previous results

EXPLAINABILITY OF RESULTS

* Explain and interpret the model
* Present feature importances and their impact level on the target

THREATS TO VALIDITY

* Document and present
  + Assumptions that may become a threat to validity
  + Uncertainties and risks
  + Data quality issues that may become a threat to validity

CONCLUSIONS

* Conclude the study and make recommendations
  + In case it is in a research phase, make it clear for business decision makers if you see a value in pursuit
  + In case it is in a development phase, transparently share the outcome and receive feedback for future improvement
* Document and communicate hypothesis testing results with the right audience

LESSONS LEARNED AND NEXT STEPS

* Document and present learned lessons
* Document and communicate
  + What can be done in the future and
  + What are the possible next steps and features for the product

# Dataset

**In-vehicle coupon recommendation Data Set**

<https://archive.ics.uci.edu/ml/datasets/in-vehicle+coupon+recommendation>

This data was collected via a survey on Amazon Mechanical Turk. The survey describes different driving scenarios including the destination, current time, weather, passenger, etc., and then ask the person whether he will accept the coupon if he is the driver.

For more information about the dataset, please refer to the paper:  
Wang, Tong, Cynthia Rudin, Finale Doshi-Velez, Yimin Liu, Erica Klampfl, and Perry MacNeille. 'A bayesian framework for learning rule sets for interpretable classification.' The Journal of Machine Learning Research 18, no. 1 (2017): 2357-2393.

<https://jmlr.org/papers/volume18/16-003/16-003.pdf>

Dataset: drivers are sent are sent coupons by their car’s mobile recommendation device. Response is whether or not driver accepts the coupon: NOT if they use the coupon

Business idea: instead of spamming coupons to all drivers, which annoys them and is distracting, better to target specific drivers under specific conditions: increase chance of them accepting and actually wanting the coupon. A way of advertising to local businesses, can be targeted according to day/time/weather.

Ford patented a system to read billboards and display ads on the infotainment display -> putting coupons instead of full advertisement as driver is nearing a restaurant is less pushy/obtrusive.

Waze already gives ads based on proximity -> something similar but with coupons

Telenav does this already -> <https://www.telenav.com/blog/why-in-car-advertising-works>

Never gives ads while vehicles are in motion: safety concern

# Resources

Telenav blog: <https://www.telenav.com/blog/why-in-car-advertising-works>

Paper data came from: <https://jmlr.org/papers/volume18/16-003/16-003.pdf>

McKinsey & Company report on monetizing car data: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/monetizing-car-data

# Paper

* Goal was to understand customers’ response to recommendations made in a in-vehicle recommender system that provides coupons for local businesses
* “Answers that the user will drive there ‘right away’ or ‘later before the coupon expires’ are labeled as ‘Y = 1’ and answers ‘no, I do not want the coupon’ are labeled as ‘Y = 0’”

# McKinsey Car Data Report

* Cars are becoming more and more connected, which means they generate more and more data. How can industry players position themselves?
* Survey: “customers are interested in data-enabled features that make mobility safer or more convenient and save them time or money”
  + Younger people more likely to want data-enabled features in cars, as well as frequent travelers
* Monetization options: data features can be “provided free of charge when customers agree to receiving advertising as part of the deal.” -> we can fit ourselves as the advertising part of the deal
  + Can be how we get data
* 3 categories of enablers:
  + In car tech like sensors, computers, data storage, location/navigation hardware
  + Infrastructural tech like 4G/5G, analytics, could computing
  + Back-end processes to connect and secure everything

From a capabilities angle, the starting point for incumbents to enter the car data monetization arena is to step up their data management capabilities in terms of:  Data preparation collection, cleansing and formatting from a multitude of relevant sources (e.g., the car, OEM Web site, social media, dealer management system)  Data analysis that applies “big data/advanced analytics” techniques to extract valuable insights from this wide and complex data landscape  Data usage and value delivery, deploying features, products, services, and recommendations to final customers and/or to business partners in order to capture the opportunity and continuously refine their offerings.

* “Industry executives concur that organizational complexity and lack of specific digital skills fundamentally hinder OEMs’ ability to innovate at the rate of nimbler high-tech players and start-ups” -> place for consulting firm to offer expertise in this domain
  + Though needs partnership with many entities: high-tech suppliers, own customers, public institutions
* 4 automobile trends behind more car data:
  + Electric cars due to climate issues and regulations
  + Advanced driver assistance systems through car connectivity
  + Shared mobility instead of private vehicles
  + Autonomous vehicles

A screenshot of a computer

Description automatically generated with medium confidence

* Compelling value prop for customers to share data/preferences: we have that with targeted coupons -> way for interested consumers to save money, tradeoff is sharing data
* Data-related use cases and business models: what we are doing
* Required technical enablers: needs more research
* Building necessary capabilities and partnerships: identify key actors in the space

## Compelling value proposition for customers

* 4 broad categories of value for customers: safety, convenience, time savings and cost reduction
  + Safety: customers willing to receive in-car advertising and share information with insurers could save money at retail POS and on insurance premiums
    - Have to develop partnership with customers
  + Graph/survey showing customers more willing to pay for time-saving use cases, many willing to have ad-based model for cost-saving
* Compared to 2015 survey, people are now “More willing to consciously grant access to their data (79 percent of globally surveyed customers, +11 percentage points vs. 2015)”
  + “More comfortable sharing personal data with apps/smartphone OS rather than for dataenabled use cases related to mobility (79 percent willingness compared to 62 percent).”
* But customers still cautious about certain data types, privacy concerns

Text

Description automatically generatedGraphical user interface, text, application

Description automatically generated

# Text Description automatically generated

# A picture containing chart Description automatically generated

* Giving customers the choice of what to share and what not to share is important: maybe build different models that rely on different tiers of private data:
  + For now, have personal demographic data -> huge privacy concern
  + Can try building a model without this personal data, using only location

## Setting business models, use cases, and monetization options to capture the opportunity

Text

Description automatically generated

* Problem for retailers: obtain customers’ clearance -> what we can do by offering coupons and several different tiers of models
* 3 main macro categories of value creation models
  + 1) Generating revenues
    - Direction monetization: not our focus
    - Tailored advertising -> what we want, targeted advertisements and promotions
    - Selling data: probably kills our product
  + 2) Reducing costs:
    - Improved customer satisfaction: the more personal data we get, the more we can tailor recommendations and predictions
  + 3) Increasing safety and security: not us
* Value of cases depends on
  + Customer adoption rate and willingness to pay
  + Value chain complexity
  + Access and ability to shape the critical control points in the value chain and technology stack requires for the use case
  + Actual ability of the organization to quickly react and leverage car data and signts
* Control points are CRITICAL to capture value creation potential

# Graphical user interface, text, application, email Description automatically generated

* We hope to capture the data capturing control points by offering a worthwhile tradeoff to drivers

## Putting the three key technical enablers in place

* 3 main enablers: in-car technologies, infrastructural technologies and back-end processes

### In-car technologies

* Ability to monetize car data hinges on the technologies development and customer acceptance of in-car tech in 8 areas:
  + Vehicle technical sensors
  + Environment sensors
    - Can sense bio-data of users: how much are they willing to share?
    - Could be used for us to make customer profiles
    - Can already do so with seat configuration in modern cars but need access to car software for that: app-system works poorly
  + High performance computing
  + Reimagined human-machine interface and customer ID
    - Voice activated software: can be used to get driver voice, allows creation of profile
    - Customer ID: NOTE: users generate data by themselves when they drive around –> can use that to get profile, use past acceptance/rejection of coupon when in certain areas to make more informed guesses
      * Music choices played on in-car software
      * Problem with confidentiality by all
      * Think of Tesla cars: have web-browser, people use that to make searches: can then tie preferences and things of interest to a specific person
  + Software platform
  + Connectivity
    - Still need constant accurate GPS and high-speed data connectivity -> push for 5G is making that possible
  + On-board data storage
  + Location/navigation
* Infrastructure technologies:
  + High speed data towers: 4G/5G
  + V2X communication
  + Smart-road infrastructure
  + Big data analytics
  + Data cloud
  + Software platforms
  + High-definition maps
* Back-end processes and operations
  + Customer privacy is important: need cybersecurity specialists
  + Retail/e-commerce opportunity made possible by car data directly involves two additional players:
    - Advertising specialists get capability to provide more targeted advertising based on enhanced car data-enable customer profiles, but this will hinge on their ability to work with OEMs to develop advertising propositions that are acceptable to drivers
      * WHAT WE ARE DOING

## Building Capabilities and Partnerships

### Organizational re(-)structuring

* “In our industry roundtables, greater freedom from rigid corporate procedures was raised often as a major advantage of start-ups (and selected high-tech players) over traditional automotive players
* Advantage of start-ups:
  + Foster digital innovation through structured incentives and allow employees to propose and experiment new ideas in working hours
  + Do not constrain early dev stages with typical conglomerate approval processes
  + Rely heavily on field experimentation rather than lengthy strategic thinking
  + Benefit from informal ideas exchanges with suppliers, partners, and even competitors
* WHOLE part about organizational structure, not very important but key point: we can position ourselves as a dynamic startup that hits the ground running and only ever picks up speed

### Strategic Partnerships

* Technology:
* Market: customers are demanding services over products
  + Growing importance of emerging markets, ex China and India, and their more progressive attitude towards digital
* Regulation:

# Value:

* Internal value: analyze survey, research approach, show that people would be willing to accept coupons while they are driving. Given the data, can try to create a profit/loss or value example
* Supervised segmentation of sorts: have to find the right conditions
* External value: how can we apply this in the real world?
  + Can advertise via coupons for any product that is related to cars: car wash, food industry, also entertainment if coupon lasts a while -> ex a Carting business
  + Why coupons and not regular advertising? Coupons give something back to customers in exchange for ad: customers less likely to be annoyed and reject ads, less backlash
  + How to get data to replicate results/predictions? Our data relies on survey data, can’t do that outside of confines of this research
    - Need way of getting demographic data
    - Distance can be from GPS data
  + Hope that demographic data matters very little

Text

Description automatically generated

# Problem & Solution

* Car data is expanding in size and scope. This provides new opportunities for data analysis that traditional car manufacturers are ill-suited to provide
  + Quote about car manufs being slow, not being agile
* There is space for start-ups to work with car manufacturers to leverage the data created by drivers and passengers
* One use case for car data is in-car advertising
* However, there are several concerns with this:
  + Ads popping up can be a safety risk
  + Advertisements can be obtrusive and obnoxious
  + To make them less annoying, can use targeted advertisement: but that requires personal data, so privacy concern
* So problem is: how do we create a service that incentivizes customers to be willing to receive advertisements in their car (whether on their phone or on the infotainment system) and also willing to share private, personal data
* Solution: give customers something in return. Make their data a currency: they provide data, we provide targeted advertisement through coupons
  + By using location and personal data, can give coupons for things that are in immediate vicinity of vehicle
  + For example, a driver who stops at a grocery store can be provided with deals/coupons that are active right now
    - Advantage for driver is clear: save money, be informed of deals/promotions
    - Advantage for grocery store: advertise certain products, put them on customers’ minds as they step into store
  + Other example: driver passing by certain restaurants on the way home, sees a coupon pop up for one of them with a discount
    - Advantage for driver: again, save money, become informed of a local restaurant
    - Advantage for restaurant: advertise the store
  + Safety concern: advertisements would only appear when car is stopped
    - Perhaps audio option?
* We are therefore at the intersection of 3 different actors:
  + Drivers/passengers: generate data, are targeted by advertisements
  + Advertisers: can be grocery stores, restaurants, car washes -> advertise themselves through deals and promotions
  + Car manufacturers: receive permission from car users to collect more personal data, gain a position into the car data market
* Where does ML come in?
  + As mentioned before, advertisements can be annoying. Irritating customers carries the risk of them not being willing to share their data
  + Goal of solution is to help customers with their data
  + So, can use ML models to target customers and serve them ads they are more likely to accept
  + Advantage for advertisers: serve targeted ads so more likely to see them have an effect
* What kind of data is used? The privacy scale
  + First tier: location/weather data
    - People are willing to share their GPS information
      * Caveat: can’t store beginning/end of journey, that’s fine we don’t need it
      * Show McKinsey graph/source
  + Second tier: preference data
    - Past history with coupons used to refine targeting
    - Can create a single customer ID for the car/phone using service, but cannot target a specific person without more information
  + Third tier: personal data (things like age, education, but also voice, music preferences, browser searches inside the car (think of Tesla infotainment system browser))
    - Can use personal data in two ways
    - First, more features for machine learning model to understand patterns, improve predictive power
      * If connected to car’s software, so working with car manuf, can use seatbelt warning to know how many passengers are in the car
    - Second, create unique customer ID for users, allows for greater preference mapping and targeting
      * Unique ID can be created using various sources: if connected to the car’s software can use seat positioning, voice recognition from drivers speaking to voice-activated software, music choices
* Different data tiers means different model tiers – able to adjust for privacy concerns
  + Baseline location/weather model can be improved on with new data, should entice users to want more