Supplemental Information

maxboonban

December 2022

1 Autonomous Freight Trucking Network

I am currently conducting research under the supervision of Dr. Pascal Van Hentenryck at the NSF AI4OPT. The goal of this project is to design an *Autonomous Transfer Hub Networks (ATHN)*, which combine L-4 self-driving trucks on middle miles while serving the first and last miles by human-operated drivers. The impact of autonomous driving is game-changing which includes no hours-of-service, improved safety, higher asset utilization, reduced empty miles to name a few. As part of the initial research, we conducted a study based on a real dedicated transportation network provided by Ryder Systems Inc. The initial study can be broken down into 2 parts:

1. Choosing Hubs in the Network
2. Produce Operational Plans (Vehicle Routing/Scheduling)

Choosing Hubs in the Network We adopted the weighted K-mean algorithm to process the preliminary hub location choice using the stops location information and the usage count data. We first used the stop latitude and longitude for the standard K-mean and then fitted the usage count of each stop as the weight of each stop. Our next step is using the centroids we picked from the weighted K-mean algorithm to find the corresponding nearest US truck stop for each centroid. We calculated the haversine distance between each centroid and each US truck stop, and chose the nearest truck stop for each centroid as our final hub location choice.

Produce Operational Plans (Vehicle Routing/Scheduling) Producing operational plan is the main part of this research project. Given the selected hubs, list of transportation tasks and fleet of vehicles, proprietary optimization models are developed to generate schedules for first/last mile trucks and middle-mile autonomous trucks. The objective of the optimization model is to create a feasible schedule that serves all the transportation task and minimizes the total relocation distance. The model is solved using three methods/models:

1. Constraint Programming
2. Column-Generation
3. Bespoke Network Flow

The results showed a potential 27-38% cost savings for a small network (southeast of United States) and 29-40% cost savings for a more extensive network. This translates into \$5.5-\$8.4 million per year due to labor cost savings and improved flexibility from autonomous trucks. This finding was a massive breakthrough in the industry, which have attracted 3 more companies to invest and fund this research further.

This leads to the next phase of the project (which I am currently tackling). This problem's motivation arises from the need of designing an *Autonomous Transfer Hub Networks (ATHN)* that further incorporate realistic constraints into the model. Since infrastructure cost is another key component in designing ATHN, we decided to incorporate hub capacity into the next phase of our project. This model will be named the *Autonomous Transfer Hub Networks with Hub Capacity (ATHNHC)*. The approach we took is similar to solving a *Service Network Design* problem. However, while most service network design models rely on discretization of time (i.e., instead of determining the exact time at which a dispatch should occur, the model determines a time interval during which a dispatch should occur), we chose to solve the problem in continuous time instead. This will be achieved through developing an iterative refinement algorithm known as the Partially Time-Expanded Networks.

Currently, I serve as the only undergraduate involved in the project, working alongside 2 Post-doctoral researcher, 1 PhD student and Dr. Pascal Van Hentenryck at Georgia Tech.

For more information about the project, please visit:

- `https://newsroom.ryder.com/news/news-details/2021/Ryder-Teams-Up-with-Georgia-Tech-for-Industrys-First-Data-Driven-Study-on-Impact-of-Autonomous-Trucking/default.aspx`
- `https://arxiv.org/abs/2201.06137`

2 Brain Tumor Detection ML Project

For the sake of patient health, the early detection of tumor growth inside the human brain is of vital importance. The earlier tumors can be caught, the easier it is for treatment plans to be created and executed that maximize the probability of successful recovery and survival. Even beyond detection, it is also important to identify the type of growth inside the brain. There are different types of brain tumors, and each type requires different treatment and can manifest themselves differently with regards to symptoms and impact on the patient. Brain tumors can be malignant or benign, and when these tumors grow, they can cause the pressure inside the skull to increase, leading to potential fatal brain damage. Magnetic Resonance Imaging (MRI) is a medical imaging technique that uses magnetic fields and radio waves to generate images of the human body. This technique can also be extended to the brain to help derive images that we can use in a machine learning setting for identification and classification. Thus, the problem being targeted here is to aid doctors in brain tumor diagnosis by creating a model for the classification of a type of brain tumor based on MRI images. Since different hospitals have their own unique MRI machines, the image quality may vary accordingly. We will address this issue by performing data augmentation techniques on our MRI image dataset. This is so that our model will be robust enough to handle different data inputs and be generalizable across different image qualities.

Various Machine Learning and Deep Learning techniques are used to classify brain tumor images into four classes. The methods are as follows:

- PCA/Dimensionality Reduction
- K-Nearest Neighbors
- Support Vector Machine
- Logistics Regression
- Random Forest
- Convolutional Neural Network
- ResNet

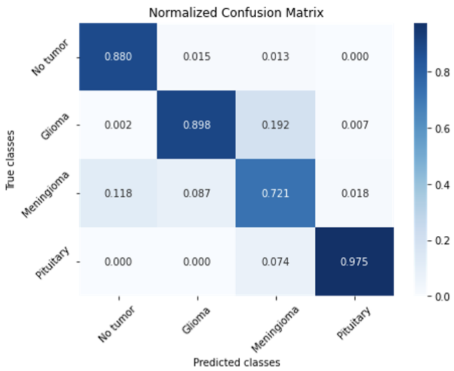
CNN The CNN passed the input through convolution, leaky ReLU activation, batch normalization, max pooling, and dropout three times before passing through 3 fully connected layers. The RMSProp optimizer was used with a decaying learning rate from 0.001 with a decay of 1e-6. The model used categorical cross-entropy loss and was trained for 5 epochs with a batch size of 32. The model had a total of 6,759,492 parameters.

ResNet The model architecture has 48 convolutional layers with 1 max pooling and 1 average pooling layer. There are shortcut connections that form identity mappings. The convolutional and identity blocks use ReLU activation. The Adam optimizer was used with a learning rate of .001 and with categorical cross-entropy loss. The model was trained for 5 epochs with a batch size of 32. The model had a total of 23,718,788 parameters.

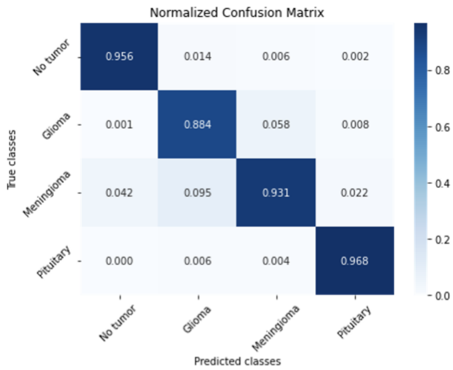
metrics / methods	Accuracy (%)		Training Time (sec)	
	Non-Augmented Dataset	Augmented Dataset	Non-Augmented Dataset	Augmented Dataset
KNN	95.3	91.2	0.1	0.5
SVM	77.0	72.2	2.2	105.7
Random Forest	91.6	89.0	0.2	1.2
Decision Tree	88.3	83.6	0.1	0.6
MLP	77.3	65.0	1.2	19.2

CNN METHODOLOGY AND RESULTS

Non-augmented data (Accuracy: 86.7%)

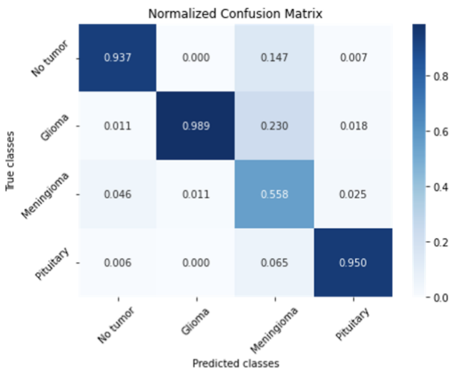


Augmented data (Accuracy: 93.6%)

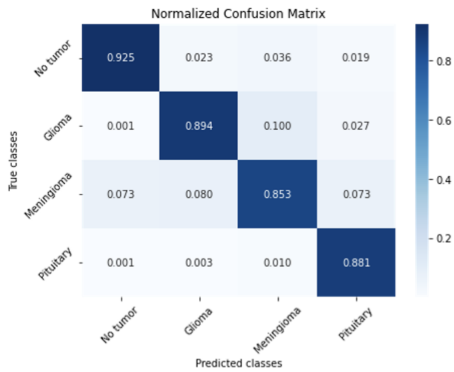


RESNET50 METHODOLOGY AND RESULTS

Non-augmented data (Accuracy: 83.8%)



Augmented data (Accuracy: 89.2%)



3 Hackathon Winning Project

Last spring, I participated at Georgia Tech Data Science Hackathon (Hackalytics). Our inspiration mainly came from the project description given to us by Georgia Tech Athletics. In the project description, there are diagrams from MLB Advanced Media, which provide a clear visual depiction of both pitches and bats. Our deliverables visually present the path that a baseball travels after being pitched and after being hit by the batter. We created a software that takes data on batting data points and suggests optimal fielder placements. We performed analysis using matplotlib, numpy and pandas and used Gurobi to perform optimization. We did a lot of data exploration through Monte Carlo simulations and settled on approaching the problem as a Mixed Integer Linear Optimization Problem. The 3D graphing library within MatPlotLib gave us many challenges, as it was not as thoroughly supported as the 2D graphing library. There were many limitations, such as the inability to draw an arc in 3D. Additionally, we had a really complex model and simplifying it to create trivial solutions on Gurobi was very difficult.

The project won the GT Athletics award and placed Top 6 overall. Additionally, we also landed a contract with the GT athletics department to continue our visualization efforts for real usage in the NCAA Division 1 regular season. We hope to integrate concepts from graph theory and Markov decision processes to improve our strategy suggestion model. We also would like to explore using reinforcement learning agents to make comparisons on natively discovered solutions to the mathematical ones.

For more information about the project, please visit:
`https://devpost.com/maxboonban?ref_content=user-portfolio&ref_feature=portfolio&ref_medium=global-nav`

4 Optimal Battery Sizing Software

This project is conducted under the supervision of Dr. Pimpa Limthongkul at Thailand's National Energy Technology Laboratory during the gap-year period I took due to COVID-19.

The increase adoption in renewables energy and decrease in price of Lithium-ion batteries is gaining momentum in the past few years. Within each micro-grid site, there is a need for a domain expert involvement in assessing the payload and designing the entire system. Hence, this incentivizes a need for a tangible tool to assist in their decision-making process. As an introductory version of the software, this model will target performing peak-shaving through the installation of photovoltaics (PV) and wind generation sources coupled with a Lithium-ion battery. Peak-shaving is defined by the intention to reduce spike in electricity consumption through the use of batteries which ultimately delays grid investment.

For more information about the project, please visit :
`http://www.ijsgce.com/uploadfile/2021/0621/20210621023720768.pdf`

5 Maritime Economics Study

This project is conducted under the supervision of Dr. Pisit Jarumaneeroj at Chulalongkorn University.

Traditional methods in measuring each countries' trade competitiveness is only able to explain either network topology or economics but not both. The goal of the project is to develop the "Container Port Connectivity Index (CPCI)" which aims to capture both economics and network topology. Then, further analysis will then move on to explain shifts in global trade pattern through the exploration of Community Structure by constructing the "Global Container-Shipping Network (GSCN)".

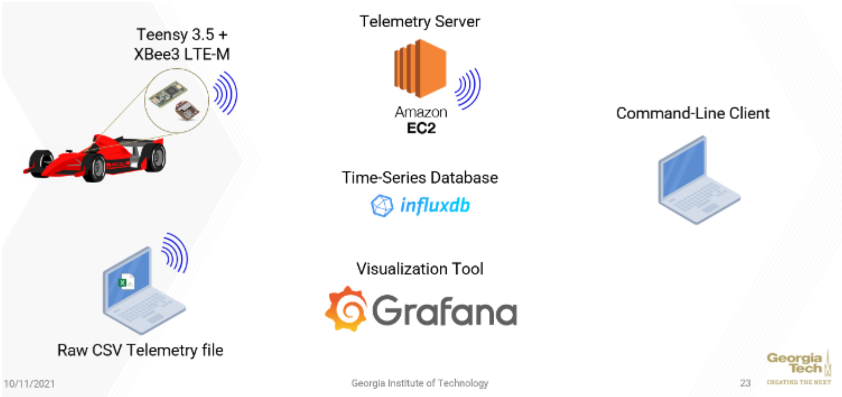
For more information, please visit (publication):
`https://link.springer.com/article/10.1057/s41278-022-00243-9`

6 Formula Electric Racing Team

HyTech Racing is Georgia Tech's student-run electric race car team. We compete in the most popular collegiate engineering competition in North America, Formula SAE. As a committed member to the club from 2018 - 2021, I oversaw overall vehicle design, managed the race car's developmental timeline, organized financial outlook, and directed the operations of the team. I was also extremely involved personally in the development of the battery pack and the telemetry system.

I will aim to explain the telemetry system project I was directly involved in. The motivation is to design an online telemetry system capable of sending data from the car to the cloud as the car is running. This solution was designed to solve two major problems with the previous Xbee PRO-based point-to-point radio design. First, range of the previous solution was limited by line of sight to the car. Second, receiving data from the radio required a cumbersome hardware setup and large amounts of team-designed code. We want to move toward a more flexible visualization platform that can easily be adapted for the addition of new sensors. As such, we chose to implement an internet-based telemetry visualization platform with an LTE modem mounted on the car.

Data is transmitted by outputting the structured data over the Teensy's serial port. The Xbee is connected to the serial port and reads the data from serial to transmit it. The Xbee is capable of running code directly on its onboard microprocessor, making it perfect for translating the messages coming from the Teensy's serial port into discrete packets to send using MQTT. The Xbee does not decode the message or interpret it in any way. It simply separates the messages according to the delimiter (0x0), appends a timestamp (acquired over the LTE network), and transmits the message to the MQTT broker. The database is installed on an AWS EC2 instance. When the car is running, the Teensy 3.5 on the main ECU monitors the can bus and relays all the messages to an LTE module, which is running Micropython code which retransmits to an MQTT server that's also running on the EC2 instance. The MQTT server pipes the data into the parser and then into the database and then into Grafana for visualization.



HyTech taught me so much about engineering design, simulation and manufacturing, and was absolutely instrumental in my education at university. It was the place where I could take everything I learned in class and apply it to a real, fully student-led project. Aside from the great success the team had at recent FSAE competitions, HyTech was one of the best parts about being at Georgia Tech also because of how much fun it was. Working late nights with other equally-passionate students was, somewhat ironically, an amazing way to unwind after a long day of lectures.

7 Honorable Mention, Thailand's Marine Department

I observed that existing oil booms didn't have enough flexibility, causing splash-over failure (phenomenon when oil leaks due the boom's inability to resonate with the movement of sea waves). My project aimed to find better materials for oil booms. I investigated other materials and their manufacturing techniques, as well as tested their properties. With help from the NCR Rubber team, I found that polyester and the twist-de-twist manufacturing method is sufficiently flexible and moves in resonance with tidal movement, significantly reducing splash-over and wrote a proposal on it. The project was awarded Honorable Mention Certificate from Thailand's Marine Department.