**Efficient Shipping**

The fourth quarter of a calendar year is a great time for retail businesses, this is when they generate most of their revenue. In the past few years shopping has transitioned from brick and mortars stores to online shopping, especially post covid. One of the most important factors of online shopping is prompt shipping. My proposal is to provide a shipping apparatus for a company like Amazon, a shipping schedule that states which route a driver should take when dropping off merchandise to customers. For example, a driver would have 10 packages to deliver in a given city and my proposal would be to provide an efficient route for delivery.

The programs attached via Jupyter use the spatial data I'll be using for the project to show EDA. The challenge of determining the most effective path between several locations will be determined via reinforcement learning. The project's objective is to use reinforcement learning to suggest a solution to package deliveries. The data is obtained from Geofabrik, a provider of OpenStreetMap data, in Protocolbuffer Binary Format (PBF). There is data accessible for a variety of means of transportation, including driving and walking - I've only mentioned the former here. Vector geographic information system data on the state of Ohio's drivable roadways make up the data I've chosen. A "way" is the term used to describe each road's digital representation. These facts will be used by me to compute the travel time, cost, and distance between sites. Gas station locations and sample addresses are the last pieces of information to be gathered.

Many ordinary computers are unable to handle the task of determining the most efficient path because P(n) grows exponentially as n grows linearly (i.e., calculating the parameters for each permutation). When compared to the daily mail distribution within each US city, this starts to become the case at a comparatively low quantity. In fact, P(75) has 2.4809140811395E+109 permutations, which is greater than the maximum estimate of atoms in the observable universe (10E+82, respectively). One approach to this issue (used by UPS) is to group the locations into groups of five and determine the fastest route for each group before grouping the groups in the same way and determining the fastest path between them. This will go on up until there are no more than five joined clusters, at which point they are all combined. This method does not find the most effective path, but it reduces the number of calculations required to locate a good route. This algorithm fails to take into consideration the possibility of more effective routes that pass outside the confines defined by the lowest-level clusters since it clusters places.

When information is available, each way is described with specifics like the number of lanes, type of access, maximum speed limit, etc. Even if a lot of this data is missing, the project is not significantly affected by this. There will need to be a maximum speed restriction for each path, and the Google Maps API is one feasible approach to fill in this information. I could also try utilizing a different city with data that is more comprehensive. For each approach, geographic spatial data is naturally available, and it is this data that will be most useful. Distributions, outliers, and correlations are probably not going to exist in this type of data in any meaningful sense because of the underlying characteristics of the data. Here, the data's objective is to present spatial information about each potential approach to influence the model's output (the most efficient route). The length and geometry columns in their whole are used to provide this spatial data. I have an idea of how the model will train. The model should work out different paths using the ‘ways’, making a decision at every intersection, and will be rewarded based on overall travel time and cost in gas.

**Questions and Answers**

1. Regarding the code, was there any problems that arose or might occur in future?

Yes, you will be tempted to download most recent versions of osmnx, NetworkX, and GeoPandas but that caused problems. I recommend you download osmnx 1.1.2, NetworkX 2.6.3, and GeoPandas 0.10.2 to avoid any problems.

1. Can you explain a ‘way’?

Yes, a ‘way’ is basically a path that is taken for a package to be delivered. This path would try to maximize speed and cost effectiveness.

1. Can you tell us more about the data?

Yes, the data for this was obtained from OpenStreetMap, I provided the link below.

OpenStreetMap or “OSM” is a free platform maintained by volunteers, it’s basically an open geographic database. OSM, is a Wikipedia for maps in which volunteers can compile information from surveys, trace it from aerial photography, and import it from other geodata sources with open licenses. Link below

<https://www.openstreetmap.org/#map=4/38.01/-95.84>

References

Dunham, J. How Data Science Can Help Shipping, Logistics And More. Forbes. Retrieved Feb 1, 2023, from [https://www.forbes.com/sites/forbestechcouncil/2022/05/26/how-data-science-](https://www.forbes.com/sites/forbestechcouncil/2022/05/26/how-data-science-can-help-shipping-logistics-and-more/?sh=12db862a5ee8)

Kuo, M. How Do Couriers Like FedEx plan their Routes. Routific. Retrieved Feb 1, 2023, from <https://blog.routific.com/blog/how-do-experts-like-fedex-plan-delivery-routes>

Bowles, N. 3 Ways Data Science Will Benefit Logistics in 2019. LateShipment. Retrieved Feb 1, 2023, from <https://www.lateshipment.com/blog/ways-data-science-will-benefit-logistics/>

Booth, G. Effective Delivery Management in a data science and AI team. LinkedIn. Retrieved Feb 3, 2023 from [https://www.linkedin.com/pulse/effective-delivery-management-data-science-](https://www.linkedin.com/pulse/effective-delivery-management-data-science-ai-team-graham-booth)

Sirota, P. Optimizing route planning for FedEx drivers: Straightaway. Medium. Retrieved Feb 3, 2023, from [https://blog.mapbox.com/optimizing-route-planning-for-fedex-drivers-](https://blog.mapbox.com/optimizing-route-planning-for-fedex-drivers-straightaway-22f18ac24a65)

Saci, S. 4 Impacting Projects to Start Your Data Science for Supply Chain Journey. Medium. Retrieved Feb 3, 2023, from <https://towardsdatascience.com/4-impacting-projects-to-start-your-data-science-for-supply-chain-journey-fe068e503c29>

Campise, K. Data Science in Logistics. Discover Data Science. Retrieved Feb 5, 2023, from <https://www.discoverdatascience.org/industries/logistics/>

Zhouyu, J. Overcoming food delivery challenges with data science. Medium. Retrieved Feb 5, 2023, from <https://medium.com/coupang-engineering/overcoming-food-delivery-challenges-with-data-science-6420cac1d59>

Eadicicco, L. How Amazon Delivers Packages in Less Than an Hour. Time. Retrieved Feb 5, 2023, from [https://time.com/4159144/amazon-prime-warehouse-new-york-city-deliveries-](https://time.com/4159144/amazon-prime-warehouse-new-york-city-deliveries-christmas/)

How Are FedEx Routes Planned? My Route Online. Retrieved Feb 5, 2023, from https://www.myrouteonline.com/blog/how-are-fedex-routes-planned