Mphasis Foundation - Mid Prep

Proposed Approach

13/11/23
Achal Jain, B22080

Problem:
Airlines often need to change their flight schedules for various reasons like seasons, new routes, or time adjustments. These changes can affect passengers who then need to be moved to other flights. We want to create a solution that can automatically figure out the best alternate flights for affected passengers based on specific rules.

Key Components:
1. **Data**: information about flight schedules, reservations, and tickets
2. **Business Rules**: rules that define how passengers should be accommodated based on factors like passenger type, loyalty, and flight details and its priority list
3. **Aim**: Find the best alternate flights for affected passengers while following these rules.

Solving this problem requires a combination of data processing, rule-based decision-making, and optimization

Dividing the problem statement into easier to solve problems/steps :-

- Data understanding and preprocessing - Understanding the structure of the given data
- Defining the business rules explicitly (**very important step**)
- Identifying what changes impact what individuals and understanding passenger preferences
- Model Design and then Optimization
- Testing and improvement
- Feedback systems for later improvements

The **first step** is quite straightforward. (ML Techniques)
The **second step** can be done manually or using graphs (Prioritization and/or sorting)(PS Says it will be given to us or use IATA data models https://airtechzone.iata.org/aidm_model/20.2/index.htm?goto=4:2:3:3:4:1:5064)
The **third step** and **fourth step** involves pattern recognition and model building. Methods we could use include

- **Algorithmic Approach** - **Dijkstra's / Bellman Ford** Algorithm (Model the problem as a graph and finding shortest distance - Could also use **topological sorts** or other modifications), **Rete Algorithm** (Popular in rule based systems), **Basic sorting** algorithms for prioritizations (This helps even out the load on system), **Genetic Algorithms** (Though I have not worked with it, it is often used in organizations for selections and approximate solutions), **Ford-Fulkerson algorithm** (could be useful for optimizing the flow of passengers through the available flights/routes - not completely

sure), **A-Search Algorithm** (good structuring method in graphs), **Simulated annealing algorithm** (Never worked with it but heard it is useful for multi parameter models but could be a little slow), **DTW** (Useful with time series)

- **ML/DL Approach - Decision Trees** (Categorize passengers), **Random Forest** (to combine multiple trees and graphs), **K Means Clustering** (Grouping passengers with similar situations), **SVMs**, **Gradient Boosting** (XGBoost for ranking and prioritization), **RNN - LSTM** (for time series related issues in flights involving sudden updates), **DQN** (Reinforcement learning approach), Agent Based Model and Generative Adversarial Networks (for creating more data for testing), **GNN**

I have mentioned some of the many algorithms we normally use in Agent Based Models, Digital Twins and Simulations (which were relevant to the given problem). The Third and fourth step would involve testing combinations of these algorithms and checking for the best approach.

What I favor most -
I primarily would like to solve the problem without any Deep Learning Technique and would like to solve it using simple rule based systems and some graph based techniques. For this I find graphs very reliable since they are easy to update.

If I have to use Neural Networks I would use GNNs as they develop on the previous method
Graphical Neural Networks: -
The way I see it we could develop majorly 4 graphs

- Passenger Flight Relations - **NODES** - Passenger and Flights, **EDGES** - Connections between passengers and the flights they are booked on, **SCOPE -** GNNs can learn patterns in passenger behavior, preferences, and interactions with flights, aiding in personalized re-accommodation recommendations while looking for similar regions in the graph (which would mean similar circumstances could be grouped)
- Flight Network Relations - **NODES** - Airports and Flights, **EDGES** - Connections between flights and Time based factors, **SCOPE -** GNNs can optimize the flow of passengers through the flight network, considering constraints and preferences by checking for connections between airports and their average delays rating of the airports
- Time Relations - **NODES** - Time Intervals and Flights, **EDGES** - Connections between flights and airport they fly to, **SCOPE -** find the regions of graphs requiring urgent attention using GNNs which would indicate the places we need to optimize. This could also help find the part which is causing most issues in the network.
- Business Rules Relation - **NODES** - Rules (IATA based models are given as graphs in their website), **EDGES** - relations between the rules, **SCOPE -** to find what rules would be triggered if a certain approach is followed

As for the **bonus things** we needed to consider in the problem statement - the above approach should be able to handle it but finding least weighted solutions

Advantages of GNNs - Robust and adaptable to change. Fast and efficient, IATA data models involve graphs, Popular in Rule based systems

Flaws with GNNs - require a large amount of data for training (could use ABM and GAN to counter this flaw)

Steps 5 and 6 : - Use ABM for testing and analysis. For self growth of the model either it could be baked in as a part of the model (neural network) or the data structure itself to analyze better methods

Some Research Papers which could be looked into before acting on some solution (They have a similar problem to the one we have to solve):-
https://airtechzone.iata.org/aidm_model/20.2/index.htm?goto=4:2:3:3:4:1:5064
https://www.sciencedirect.com/science/article/pii/S0965856401000398
https://www.sciencedirect.com/science/article/pii/S0360835219304747
https://pubsonline.informs.org/doi/abs/10.1287/trsc.1090.0269
https://sciencedirect.com/science/article/pii/S0968090X21001480
https://www.researchgate.net/publication/342851755_Optimization_of_Flight_Rescheduling_Problem_under_Carbon_Tax
https://onlinelibrary.wiley.com/doi/10.1111/itor.13079
https://onlinelibrary.wiley.com/doi/10.1111/itor.13396
https://enac.hal.science/hal-03701665/document
https://www.researchgate.net/publication/339330941_Multiobjective_Evolutionary_Scheduling_and_Rescheduling_of_Integrated_Aircraft_Routing_and_Crew_Pairing_Problems