Towards the Next Generation Airline Revenue Management: A Deep Reinforcement Learning Approach to Seat Inventory Control and Overbooking

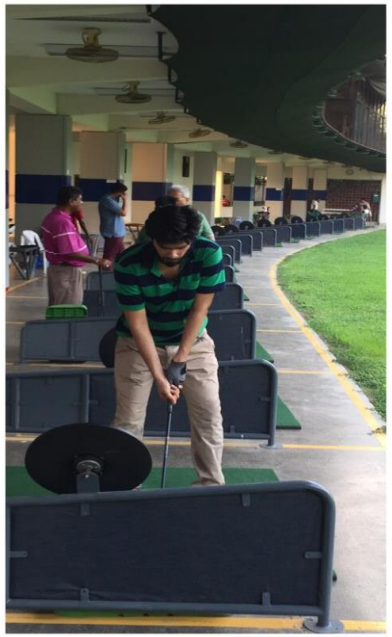
By **Syed Shihab**, Caleb Logemann, Deepak-George Thomas, and Peng Wei

Intelligent Aerospace Systems Lab

Iowa State University

Presenter's bio

- Born in Jashore, Bangladesh
- Bachelor's in EEE, AIUB, Bangladesh, 2012
- PhD in Aerospace Eng., ISU, Fall 2019 (expected)
- Research interests: operations research, value-based engineering design, decision making under uncertainty, ML/AI/RL
- Fun Fact: I was ambushed by bears two summers back on a cycling trail in Aspen, Colorado

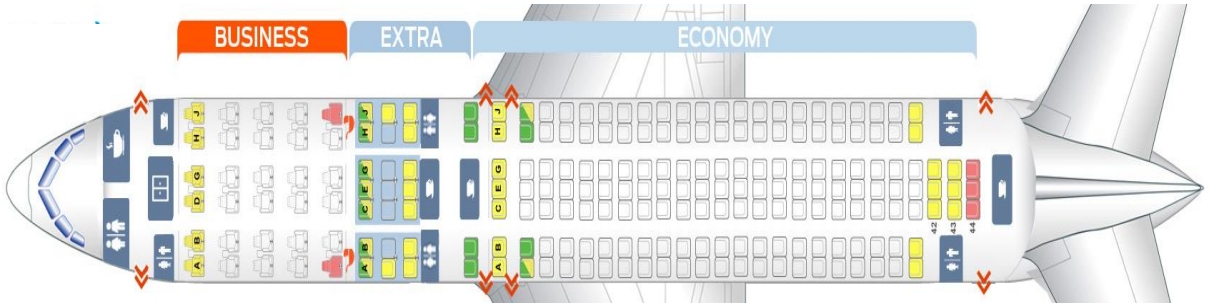


Outline

1. Introduction to airline Revenue Management (RM), Reinforcement Learning (RL) and deep RL
2. Motivation behind using deep RL approach
3. RM problem formulation
4. Experiments and Results
5. Conclusion and future work

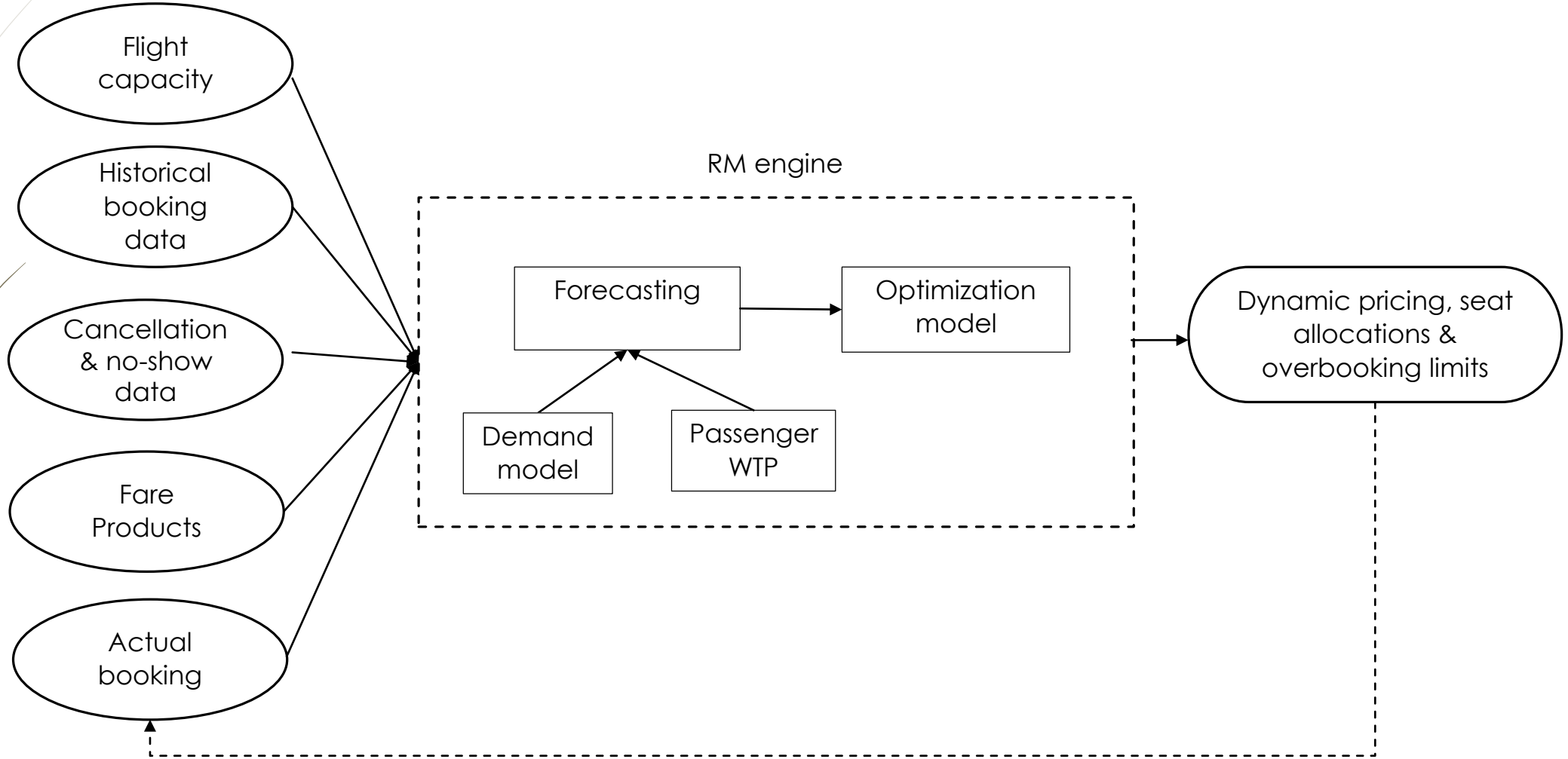
Revenue Management (RM)

- Objective: maximize revenue from all scheduled flight departures
- Decision making problems:
  - Pricing
  - Seat inventory control
  - Overbooking (OB)



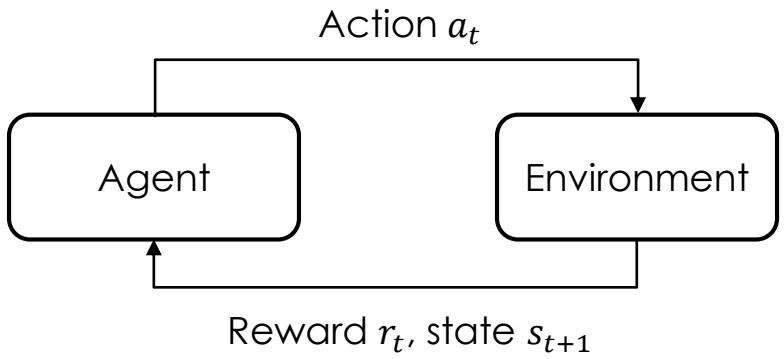
1. https://pointmetotheplane.boardingarea.com/2018/09/21/airline-change-fees-might-and-should-go-the-way-of-the-dodo/
2. http://theflight.info/seat-map-boeing-767-300-american-airlines-best-seats-in-the-plane/

RM system components



Reinforcement Learning

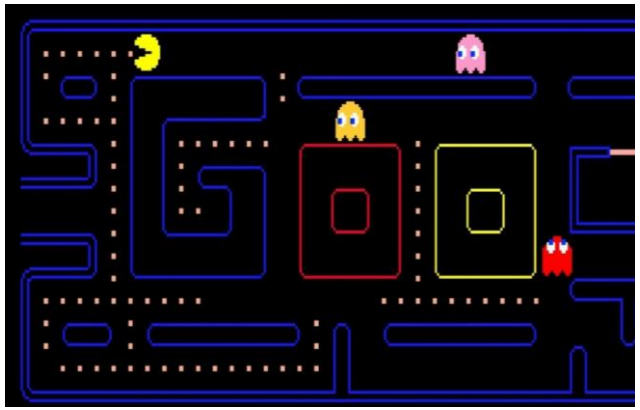
Idea: Agent learns by interacting with environment



Goal: find the optimal policy $\pi^* : S \rightarrow A$

$$\pi^* = \underset{\pi}{\operatorname{argmax}}\, \mathbf{E}\left[\gamma^t \sum_{t=0}^{T} R(s_t, a_t) | \pi\right]$$

Application domain: Sequential decision-making problems in stochastic environments



1. https://www.wired.com/2010/05/pac-man-google/
2. https://www.semanticscholar.org/paper/End-to-End-Training-of-Deep-Visuomotor-Policies-Levine-Finn/
3. https://www.offerzen.com/blog/how-to-build-a-content-based-recommender-system-for-your-product

Reinforcement Learning

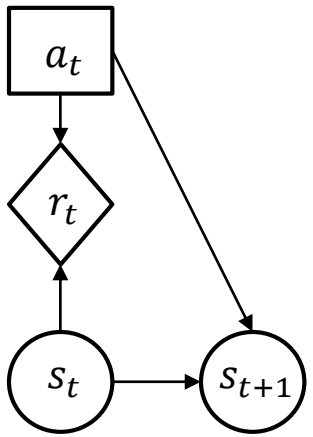
- Modeling approach: Markov Decision Process (MDP)
- MDP components: states, actions, transition function and reward function
- Value of a state $U^*(s)$

$$U^*(s) = \max_a (R(s,a) + \gamma \sum_{s'} T(s'|s,a)\, U^*(s'))$$

- State-action values $Q^*(s,a)$

$$Q^*(s,a) = R(s,a) + \gamma \sum_{s'} T(s'|s,a) \max_a Q^*(s,a))$$

- Agent does not have any knowledge of model dynamics



Q-learning

- Popular model-free RL algorithm
- Directly estimates $Q^*(s,a)$ from experience
- Incremental update rule:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left(r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right)$$

- Balancing exploration and exploitation
  - Exploit: Best action = $\max_a Q^*(s,a))$
  - Explore: Random action

Deep RL

- Deep neural network is used for approximating the $Q^*(s, a)$ values
- Domain knowledge is typically given as input to facilitate learning
- Useful for problems with large state spaces
  - Not possible to visit all states
  - Not possible to tabulate all values
- Agent can generalize from limited experience to states that have not yet been visited

Motivation

- Why use Reinforcement Learning (RL) in airline RM?
  - Airline RM problem can be modeled as a MDP
  - Potential alternative to existing RM systems?
    - No need for complex revenue model
    - After initialization, agent's performance does not depend on demand model
    - Room for exploration
- Why use deep RL?
  - Generalization for large continuous state space
  - Approximation of highly nonlinear state-action value function by incorporating domain knowledge

RM Problem Statement

- Single Origin-Destination market
- Three fare classes: high ($300), middle ($200) and low ($100)
- Air travel market characteristics:
  - Booking window = 365 days
  - Arrival of passengers is modeled as a Poisson process
  - Cancellation probability
  - Passenger cancellation times follow a uniform distribution
- Bumping Cost (BC) = Fare reimbursement + goodwill cost
- Flight capacity = 80
- Avg. # booking requests/flight = 100
- Goal: maximize revenue from each flight
- Task: optimally allocate seats and set overbooking levels for each fare class

MDP Formulation

- State representation

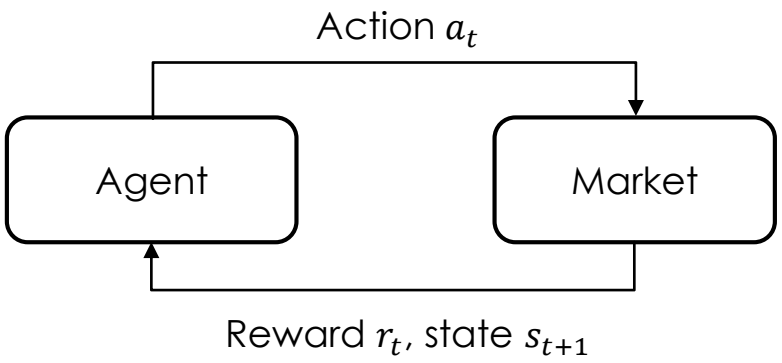
$$S = (T, b_1, b_2, b_3, t)$$

- Actions: accept, deny

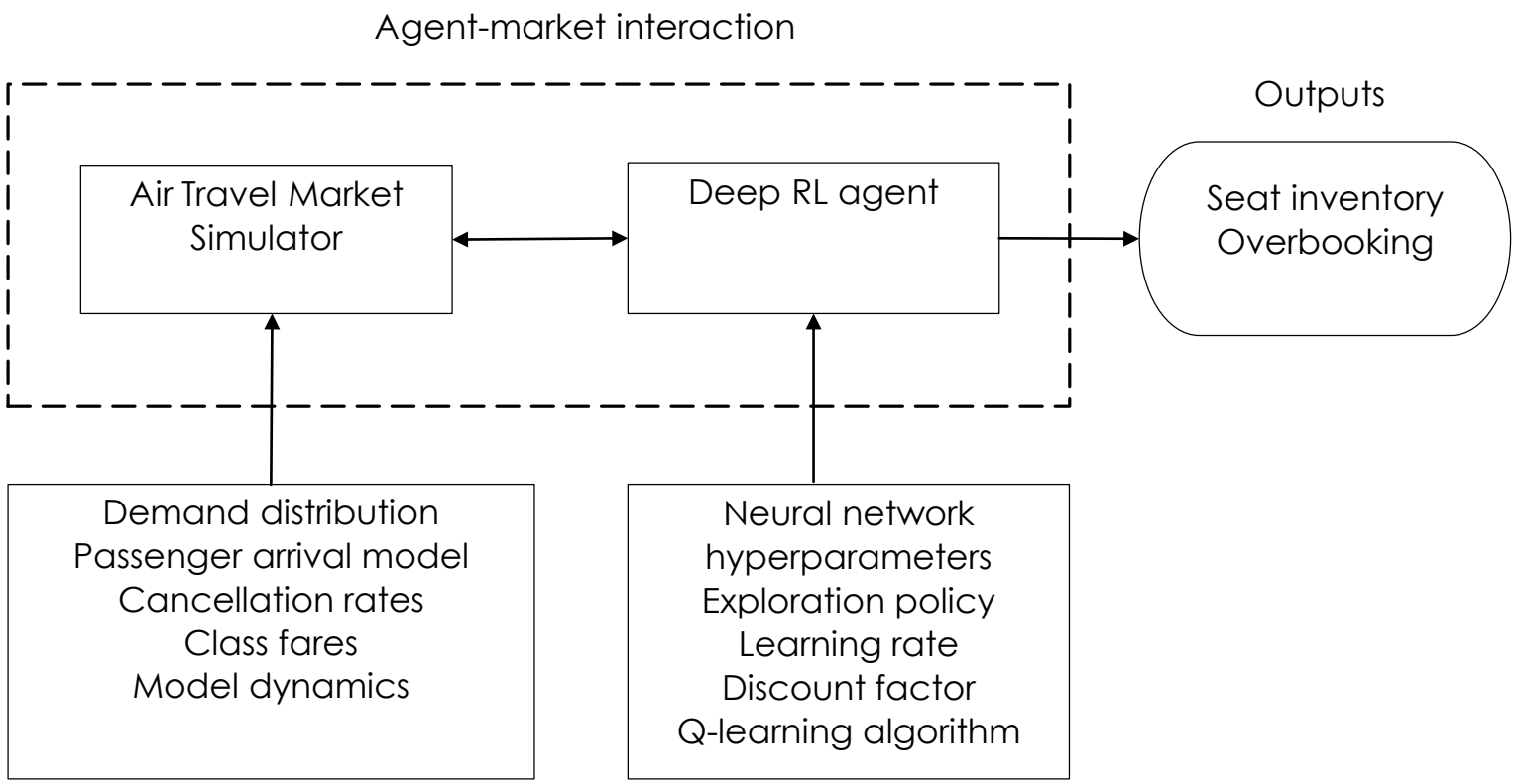
$$A = \{a_{+1}, a_{-1}\}$$

- Model dynamics
  - When a customer requests booking, agent needs to take an action
  - If booking request is accepted, increment corresponding booking class variable
  - If a booking gets cancelled, decrement corresponding booking class variable
  - Episode ends at t=0 (flight departure)
- Reward function

$$R = \begin{cases} f_T & Action = a_{+1} \\ 0 & Action = a_{-1} \\ -f_T & \text{Occurrence of cancellation} \\ -BC & \text{at t} = 0 \end{cases}$$

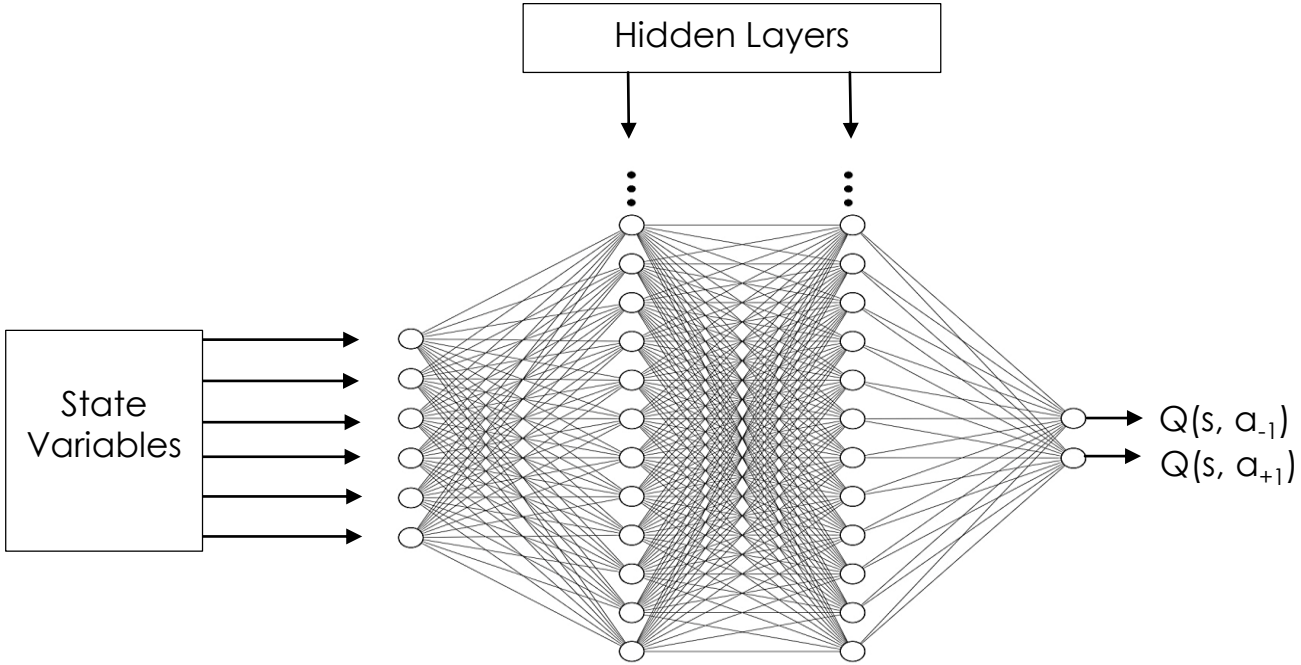


Solution Method



Neural Network Configuration

- 2 hidden layers
- 128 nodes in each hidden layer
- Relu activated layers
- Exploration strategy: linear annealed epsilon-greedy policy
- Implemented using Keras and Keras-rl packages of Python



Experimental Results

Fare class distribution	Mean booking requests
1	[10,30,60]
2	[60,30,10]
3	[33,33,34]

- BC factor = 2
- Theoretical optimal revenue
  - Filling the aircraft to capacity starting with high fare passengers to lower fare ones without any bumping

Cancellation rate (%)	Fare class distribution	Avg. optimal revenue (%)	Avg. acceptance (%)	Avg. load factor (%)
0	1	97.628	77.67	96.365
0	2	93.651	81.715	101.738
0	3	97.046	77.434	95.092
10	1	95.564	88.014	97.995
10	2	93.661	89.418	100.608
10	3	94.855	78.707	88.787
20	1	96.08	96.09	95.984
20	2	89.868	81.949	81.909
20	3	94.637	93.407	93.317

Experimental Results

Fare class distribution	Mean booking requests
1	[10,30,60]
2	[60,30,10]
3	[33,33,34]

- BC factor = 2
- Theoretical optimal revenue
  - Filling the aircraft to capacity starting with high fare passengers to lower fare ones without any bumping
- In most cases, agent achieved above 93 % optimal revenue

Cancellation rate (%)	Fare class distribution	Avg. optimal revenue (%)	Avg. acceptance (%)	Avg. load factor (%)
0	1	97.628	77.67	96.365
0	2	93.651	81.715	101.738
0	3	97.046	77.434	95.092
10	1	95.564	88.014	97.995
10	2	93.661	89.418	100.608
10	3	94.855	78.707	88.787
20	1	96.08	96.09	95.984
20	2	89.868	81.949	81.909
20	3	94.637	93.407	93.317

Experimental Results

Fare class distribution	Mean booking requests
1	[10,30,60]
2	[60,30,10]
3	[33,33,34]

- BC factor = 2
- Theoretical optimal revenue
  - Filling the aircraft to capacity starting with high fare passengers to lower fare ones without any bumping
- In most cases, agent achieved above 93 % optimal revenue

Cancellation rate (%)	Fare class distribution	Avg. optimal revenue (%)	Avg. acceptance (%)	Avg. load factor (%)
0	1	97.628	77.67	96.365
0	2	93.651	81.715	101.738
0	3	97.046	77.434	95.092
10	1	95.564	88.014	97.995
10	2	93.661	89.418	100.608
10	3	94.855	78.707	88.787
20	1	96.08	96.09	95.984
20	2	89.868	81.949	81.909
20	3	94.637	93.407	93.317

Cancellation rate (%)	Average optimal acceptance (%)
0	80
10	88
20	100

Experimental Results

Fare class distribution	Mean booking requests
1	[10,30,60]
2	[60,30,10]
3	[33,33,34]

- BC factor = 2
- Theoretical optimal revenue
  - Filling the aircraft to capacity starting with high fare passengers to lower fare ones without any bumping
- In most cases, agent achieved above 93 % optimal revenue

Cancellation rate (%)	Fare class distribution	Avg. optimal revenue (%)	Avg. acceptance (%)	Avg. load factor (%)
0	1	97.628	77.67	96.365
0	2	93.651	81.715	101.738
0	3	97.046	77.434	95.092
10	1	95.564	88.014	97.995
10	2	93.661	89.418	100.608
10	3	94.855	78.707	88.787
20	1	96.08	96.09	95.984
20	2	89.868	81.949	81.909
20	3	94.637	93.407	93.317

Cancellation rate (%)	Average optimal acceptance (%)
0	80
10	88
20	100

Experimental Results

Fare class distribution	Mean booking requests
1	[10,30,60]
2	[60,30,10]
3	[33,33,34]

- BC factor = 2
- Theoretical optimal revenue
  - Filling the aircraft to capacity starting with high fare passengers to lower fare ones without any bumping
- In most cases, agent achieved above 93 % optimal revenue

Cancellation rate (%)	Fare class distribution	Avg. optimal revenue (%)	Avg. acceptance (%)	Avg. load factor (%)
0	1	97.628	77.67	96.365
0	2	93.651	81.715	101.738
0	3	97.046	77.434	95.092
10	1	95.564	88.014	97.995
10	2	93.661	89.418	100.608
10	3	94.855	78.707	88.787
20	1	96.08	96.09	95.984
20	2	89.868	81.949	81.909
20	3	94.637	93.407	93.317

Cancellation rate (%)	Average optimal acceptance (%)
0	80
10	88
20	100

Experimental Results

Fare class distribution	Mean booking requests
1	[10,30,60]
2	[60,30,10]
3	[33,33,34]

- BC factor = 2
- Theoretical optimal revenue
  - Filling the aircraft to capacity starting with high fare passengers to lower fare ones without any bumping
- In most cases, agent achieved above 93 % optimal revenue

Cancellation rate (%)	Fare class distribution	Avg. optimal revenue (%)	Avg. acceptance (%)	Avg. load factor (%)
0	1	97.628	77.67	96.365
0	2	93.651	81.715	101.738
0	3	97.046	77.434	95.092
10	1	95.564	88.014	97.995
10	2	93.661	89.418	100.608
10	3	94.855	78.707	88.787
20	1	96.08	96.09	95.984
20	2	89.868	81.949	81.909
20	3	94.637	93.407	93.317

Cancellation rate (%)	Average optimal acceptance (%)
0	80
10	88
20	100

Experimental Results

Fare class distribution	Mean booking requests
1	[10,30,60]
2	[60,30,10]
3	[33,33,34]

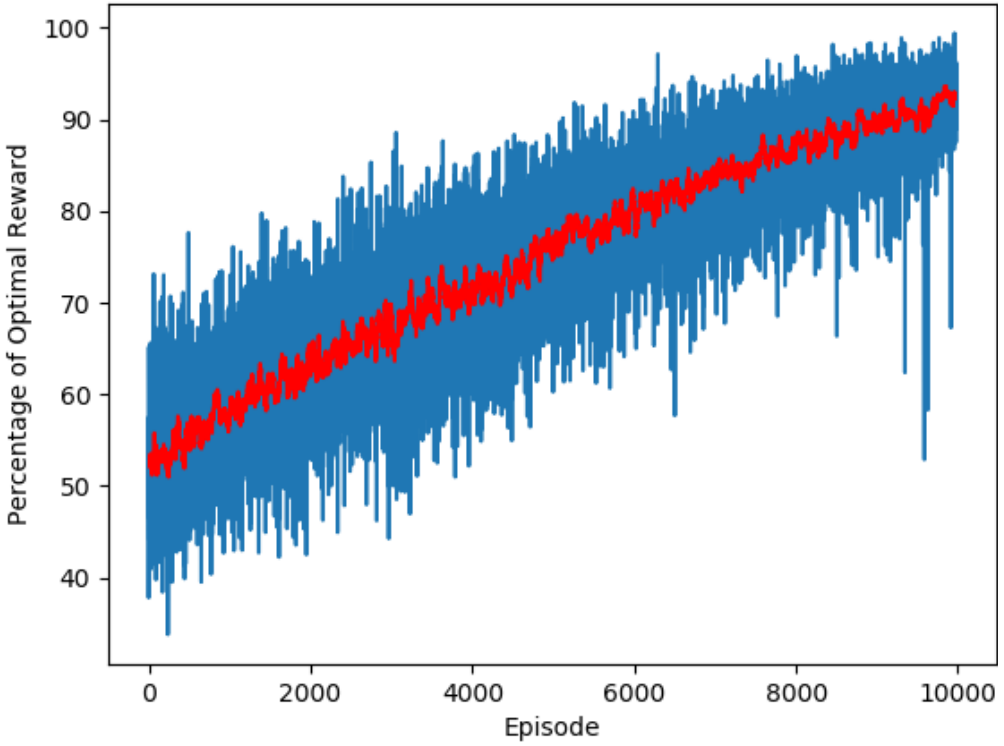
- BC factor = 2
- Theoretical optimal revenue
  - Filling the aircraft to capacity starting with high fare passengers to lower fare ones without any bumping
- In most cases, agent achieved above 93 % optimal revenue

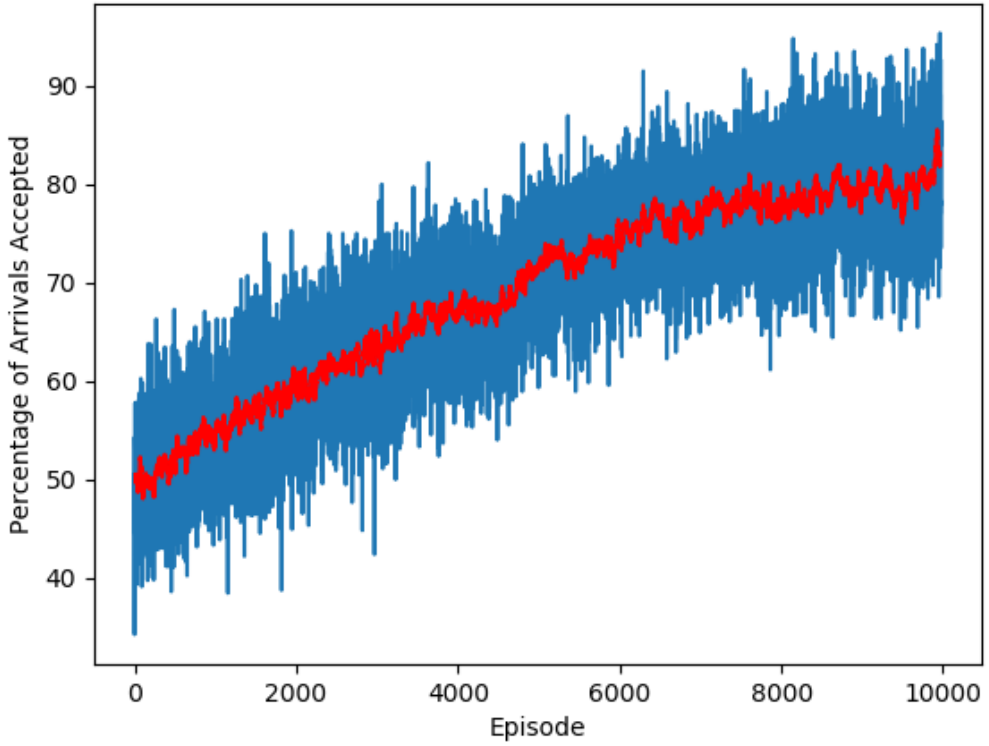
Cancellation rate (%)	Fare class distribution	Avg. optimal revenue (%)	Avg. acceptance (%)	Avg. load factor (%)
0	1	97.628	77.67	96.365
0	2	93.651	81.715	101.738
0	3	97.046	77.434	95.092
10	1	95.564	88.014	97.995
10	2	93.661	89.418	100.608
10	3	94.855	78.707	88.787
20	1	96.08	96.09	95.984
20	2	89.868	81.949	81.909
20	3	94.637	93.407	93.317

Cancellation rate (%)	Average optimal acceptance (%)
0	80
10	88
20	100

Training plots for Test Case 6

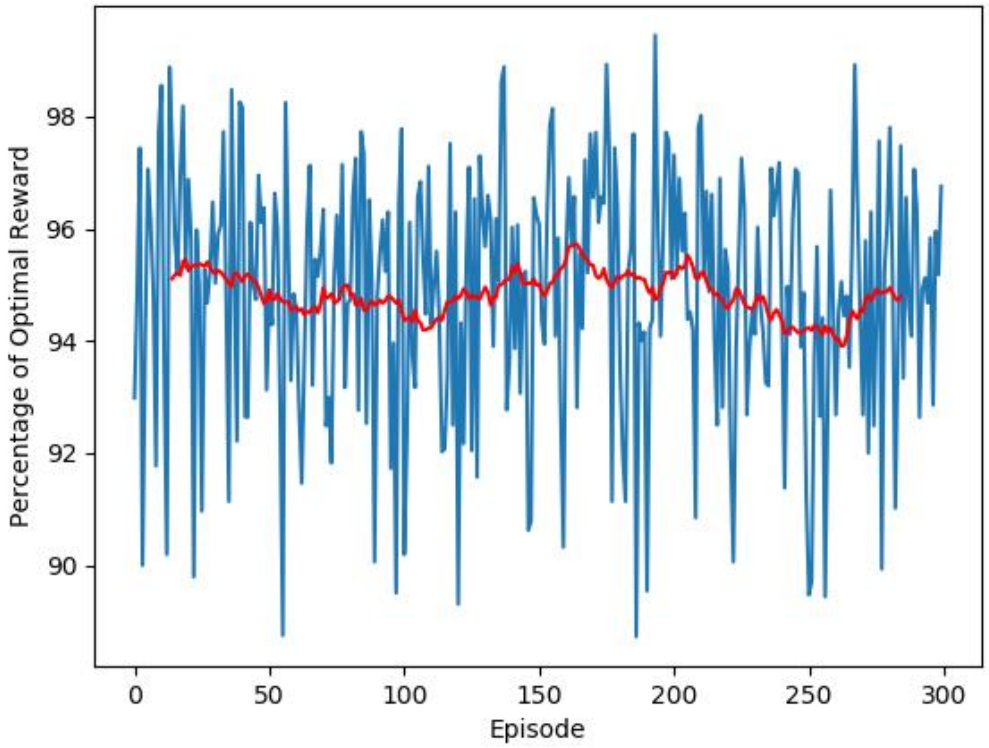
- As the training progresses, agent learns how to control seat inventory and set overbooking limits

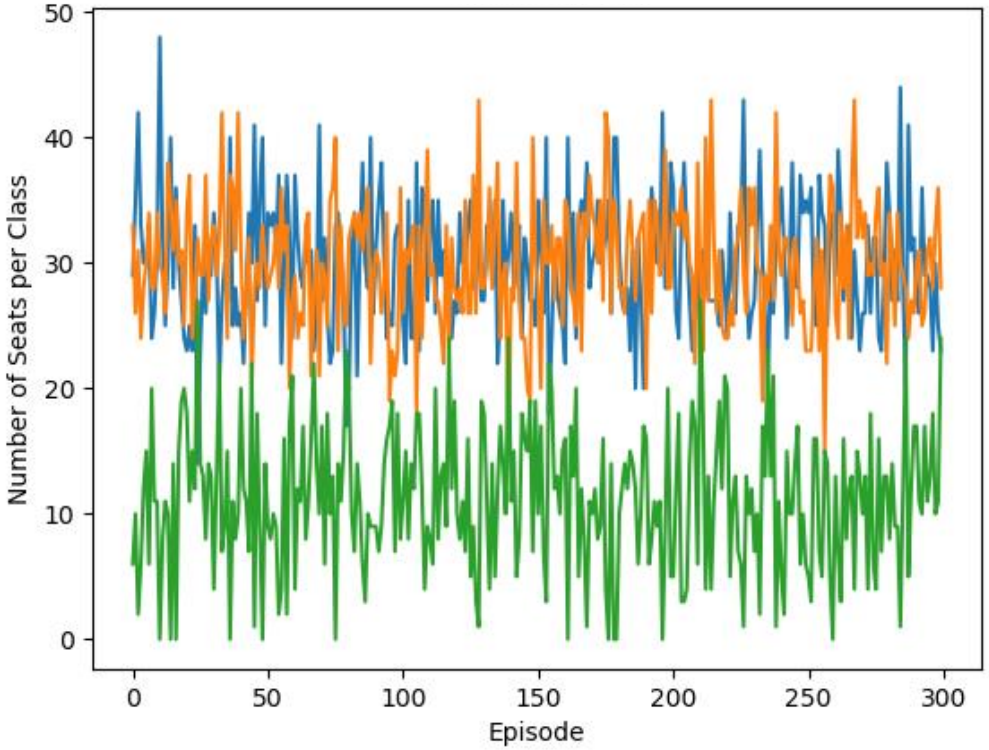




Test plots for Test Case 6

- Mean booking requests for the high, middle and low fare classes: 33, 33, and 34
- On average, agent allocates 33 seats for the high and middle fare classes, and the remaining 13 seats for the low fare class





Conclusion

- In most cases, the deep RL agent achieved above 93 % optimal revenue on average
- Performance of the agent improves with each episode
  - Theoretically, given sufficient number of episodes, agent will learn the optimal policy
- Performance of the agent is sensitive to values of hyperparameters
- Suboptimal performance when bumping cost is low
  - Higher exploration should improve results
- Next steps
  - Dynamic pricing: include additional state variables for prices of airline's and competitors' fare products
  - Benchmarking the performance of the algorithm with the EMSR method
  - Network of flight legs
  - Simulating demand disruptions

Questions and Comments?

Syed A.M. Shihab
shihab@iastate.edu