**Paper outline**

**Working Title:** Capacitated Vehicle Routing Problem (CVRP)

**Authors:** Amandeep Rathee, Ronald Tinashe Nhondova, Tzu-Chun Hsieh

**Hypothesis/Goal/Aim:** Construct a new RL algorithm that achieves higher performance/uses less computational resources/takes less training time compared to existing methods that solves CVRP.

**Objectives with (short, one-sentence) justification:** (Should serve to revisit/update your story board)

**Audience:**

**Style:** Academic paper

**Introduction**

The goal in the CVRP is to find optimal routes for multiple vehicles visiting a set of locations. This is a comm When we have a single vehicle, the problem reduces to a traveling salesman problem (TSP). The optimal solution could mean either minimizing the time, distance, or a combination of both while visiting all the locations in the route. Each location is visited only once, by only one vehicle, and each vehicle has a limited capacity. The problem also could have a time constraint where each location must be visited within a certain amount of time.

There are various ways to solve CVRP. A naive way is to do a combinatorial search over the search space and choose the path that minimizes the objective (distance or time). However, the problem grows exponentially as the number of nodes and vehicles increase and quickly becomes infeasible to be solved by this naive approach. We will use reinforcement learning (RL) to try to solve CVRP since other conventional approaches such as dynamic programming have struggled to find a solution when the size of the problem gets realistic.

**Previous work**

There have been many studies using various kinds of methods to solve CVRP problems. Salimans et al. (2017) used highly parallelizable Evolution Strategies (ES) to achieve competitive performance to RL algorithms with less wallclock time. An attention model trained with REINFORCE applied by Kool et al., 2019 is suitable for multiple routing problems and as effective as problem-specific approaches. Lu et al., 2020 introduced a learning-based algorithm for solving CVRP that iteratively improves or perturbs the initialized feasible solution to explore better solutions.

**Our approach**

**Methodology**

Explain the model we will construct in detail.

**Experiments and results**

Compare our model results to existing models on performance metrics to be determined.

**Conclusion/Discussion**

(Place holder)

**References**

* Tim Salimans, Jonathan Ho, Xi Chen, Szymon Sidor, Hya Sutskever. Evolution strategies as a scalable alternative to reinforcement learning, 2017.
* Wouter Kool, Herke van Hoof, Max Welling. Attention, learn to solve routing problems!, 2019.
* Hao Lu, Xingwen Zhang \* & Shuang Yang. A learning-based iterative method for solving vehicle routing problems, 2020.