# Estado da Arte – Machine Learning

Problemática em causa: Aprendizagem Automática (AA) no problema dos transportes públicos terrestres (autocarros, metro, comboio, etc.).

Para que possamos usufruir de todas as qualidades de machine learning temos que antes garantir a qualidade dos dados utilizados. A tecnologia de smart cards permite uma recolha de dados fidedigna e eficaz tendo sido o método de recolha utilizado na pesquisa conduzida por Florian Toqué (Florian Toqué, 2017).

Os dados recolhidos são utilizados para prever a necessidade da deslocação dos utilizadores, é notável referir que existem diferenças entre o estudo da Origem-Destino e da fluidez de viagem. O estudo em causa foca-se na fluidez contrastando a importância da previsão a curto prazo, mas não esquecendo a possibilidade do planeamento a longo prazo, que seria possível com o aproveitamento de machine learning.

Existem vários exemplos que comprovam a utilidade de *machine* *learning* no desenvolvimento e melhoramento de horários nos autocarros nas grandes metrópoles, inclusive, um estudo feito por autores portugueses baseado na cidade do Porto (João Mendes-Moreira, 2015). Foram analisados dados recolhidos pelo *Automatic Vehicle Location* (AVL), equipamentos de comunicação baseados em localização GPS que enviam o *status* com a atual localização dos vários autocarros da frota.

O método de estudo utilizado por João Mendes-Moreira consiste em formar vários *clusters* (viagens, etc.) com informação seletiva, que são estudados de forma a obter vários valores de interesse, sendo o mais importante para a nossa pesquisa o TTP, ou *travel* *time prediction,* que consiste na duração total estimada das viagens. Tal como o que foi desenvolvido ao longo do projeto, devemos começar por criar um perfil do dia que contem, para uma determinada rota, informação sobre as viagens. A aplicação deste tipo de estudo permite uma análise mais detalhada e aprofundada.

Quanto à metodologia que envolve o *machine learning* propriamente dito, implica a aplicação de técnicas como o ‘*consensus clustering’* e *‘rule induction’* para descobrir informação relevante em grandes quantidades de dados.

Acreditamos que a pesquisa estudada seja muito apropriada para o trabalho que está a ser desenvolvido na unidade curricular de LAPR5, uma vez que refere diretamente a utilização de *machine learning* no planeamento de horários de autocarros, para além de que menciona a utilização de regras nas medidas consideradas. Apesar de ser um artigo publicado em 2015, continua a ser relevante.

*Machine Learning* pode ser utilizado também para a estimativa e previsão das exigências de viagens como se lê no artigo de Florian Toqué. Este explora de forma mais aprofundada a utilização de métodos de machine learning, tendo atenção às vantagens a curto e longo prazo. Utilizando as *digital footprints* existentes como dados a ser estudados é possível criar ferramentas que permitem prever a fluidez de passageiros nas cidades grandes e aumentar os níveis dos serviços e agendamento de transportes.

As vantagens da utilização de machine learning no desenvolvimento das soluções pretendidas é aparente, permite que sejam processadas múltiplas rotas em simultâneo e garante um maior aprofundamento de dados. Uma vez que é capaz de processar dados a um ritmo maior, seria possível diminuir os custos tanto para a empresa responsável por planear as rotas, como também, eventualmente, aos clientes, com o intuito de manter um preço competitivo.

# Referencias

João Mendes-Moreira, L. M.-M. (2015). Validating the coverage of bus schedules: A Machine Learning approach. *Information Sciences*, 299-313.