**PROPOSAL** with Methodology

We would like to study “**How air traffic affected the spread of Covid**-19” from an originating country to the rest of the world. Since the outbreak of the pandemic, popular imagination has held the belief that: “had not been air traffic, Covid-19 wouldn’t have become a pandemic.” However, there has been little formal analysis demonstrating this claim. Our study aims to contribute to this less explored area, yet relevant, by examining whether countries with more arriving international travellers experienced higher COVID-19 infection rates than those that have less.

While examining a range of similar research articles, we have focused: firstly, on papers dealing with human mobility; secondly, on studies modelling the spread of Covid; thirdly, on papers discussing inflight transmission; and finally, on the few papers investigating how Air traffic was a factor for Covid-19 spread. For instance, findings from articles on inflight transmission have helped us narrow down the scope of our studies; because, it appears that only journeys of long hours have an impact on the spread of Covid-19. Consequently, our focus will be exclusively on countries; domestic air travel doesn’t appear to be an important factor.

Our hypothesis is that international Air traffic was the main factor for Covid-19 spread; making it reach global proportions. We will be investigating the spatiotemporal spread of COVID-19 between the originating country and the affected one in three phases: before travel restrictions, during travel restrictions, when special flights were put in place, and after the restrictions. Our goal is to propose a model that considers China, the country where the Alpha variant originated, as the predictor and the other countries as a dependent variable. Subsequently, repeating our experimentation by considering India as the country of origin for the Delta variant. We will successively compare the results of **Statistical Modelling from existing studies against two machine learning algorithms** of our choice namely, **Prophet algorithm** of Facebook **and Neural network**. Many studies have shown that Stochastic models hold good when it comes to analysing the spread of a virus. Machine Learning algorithms having a stochastic nature and being data driven, we believe they will yield promising results in terms of accuracy.

The spread of covid through air traffic is from a point A to another B, that is we will not consider social influences or the like as factors. Thus, we will rather directly concentrate on examining the correlation between Air Traffic and Covid-19 spread between countries A and B.

We want our datasets to be disparate. The reason being “t**he performance of the same prediction approach can differ widely across datasets”.** We are developing strong argumentation vis-a-vis datasets selection. Furthermore, we will be efficiently performing feature engineering as well as tuning hyper parameters in order to yield better results.Also while we are concentrating on scientific papers, we are keeping track of semi formal articles on Machine Learning which have dealt with this problem in order to better our intuition by looking at the problem from other angles.