**CHAPTER 1**

**INTRODUCTION**

## 1.1 GENERAL

Link prediction is the problem of predicting the formation of new edges on a given network. It is a fundamental problem that applies to networking in numerous contexts, including the Internet, the web, and online social networks. The problem can be stated as follows - given a snapshot of a dynamic social network, represented as a directed graph of nodes and edges (where nodes represent users and edges represent follower - followee relations) is it possible to predict which new relations (links) are likely to be formed in the future.

Numerous edges and vertices are added and/or removed throughout time as a consequence of changing individual connections. Consequently, social networks become very dynamic and complex. Identifying the mechanisms by which they evolve is a fundamental question that is still not well understood, and it forms the motivation for our work here.

## 1.2 NEED FOR THE STUDY

From the link prediction research perspective, many supervised machines learning (ML) classification techniques may be used to address the link prediction problem. Multiple studies have shown that this method delivers good results; nevertheless, choosing the collection of features (variables) to train classifiers remains a key challenge. The issue of link prediction in dynamic or time-varying networks involves two key challenges: precision and efficiency.

Additionally, because similarity-based link prediction algorithms were first developed for static networks, it is crucial to choose which similarity-based approaches and machine learning algorithms to use when applying them to dynamic networks.We conduct extensive experiments with different machine learning algorithms to address this issue as well.

## 1.3 OBJECTIVE OF THE STUDY

We aim to provide a solution to the issue of link prediction in dynamic networks using supervised learning machine learning algorithm and choosing the best model out of it. We attempted to solve the link prediction problem for a dataset gathered from Twitch - an online video streaming platform.

## 1.4 MOTIVATION FOR USING CLASSIFICATION ML ALGORITHMS

* As a technical problem, the efficacy of link prediction is generally not well understood. Today, link prediction algorithms are the basis for social recommendations in a wide range of social networks and applications, ranging from Facebook and Pinterest and Q&A sites like Quora. The success of these sites and the sheer volume of prior literature lead many to believe the problem is well addressed. Only evidence to the contrary comes from failed recommendations.
* Link prediction is frequently used in social networks to suggest friends to users. It has also been used to predict criminal associations.
* In biology, link prediction has been used to predict interactions between proteins in protein-protein interaction networks. Link prediction has also been used to infer interactions between drugs and targets using link prediction.
* Link can be used to improve marketing strategies, in particular viral marketing.