Image Popularity Prediction (IPP) Over Time Using Machine Learning Techniques

Author

Amna Shahid

319726

Supervisor

Dr. Muhammad Usman Akram

DEPARTMENT OF COMPUTER ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD

July, 2021

Image Popularity Prediction (IPP) Over Time Using Machine Learning Techniques

Author

Amna Shahid

319726

A thesis submitted in partial fulfillment of the requirements for the degree of

MS Computer Engineering

Thesis Supervisor

Dr. Muhammad Usman Akram

Thesis Supervisor's Signature: ____________________________________

DEPARTMENT OF COMPUTER ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY, ISLAMABAD

JULY, 2021

Declaration

I certify that this research work titled "*Image Popularity Prediction (IPP) Over Time Using Machine Learning Techniques"* is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources it has been properly acknowledged / referred.

Signature of Student

Amna Shahid

319726

Language Correctness Certificate

This thesis has been read by an English expert and is free of typing, syntax, semantic, grammatical and spelling mistakes. Thesis is also according to the format given by the university.

Signature of Student

Amna Shahid

319726

Signature of Supervisor

Dr.Muhammad Usman Akram

Copyright Statement

Acknowledgements

All praise and glory to Almighty Allah (the most glorified, the highest) who gave me the courage, patience, knowledge and ability to carry out this work and to persevere and complete it satisfactorily. Undoubtedly, HE eased my way and without HIS blessings I can achieve nothing.

I would like to express my sincere gratitude to my advisor Dr. Muhammad Usman Akram for boosting my morale and for his continual assistance, motivation, dedication and invaluable guidance in my quest for knowledge. I am blessed to have such a co-operative advisor and kind mentor for my research.

Along with my advisor, I would like to acknowledge my entire thesis committee: Dr. Sajid Gul Khawaja and Dr. Arslan Shaukat for their cooperation and prudent suggestions.

My acknowledgement would be incomplete without thanking the biggest source of my strength, my family. I am profusely thankful to my beloved parents who raised me when I was not capable of walking and continued to support me throughout in every department of my life and my loving sisters who were with me through my thick and thin.

Finally, I would like to express my gratitude to all my friends for always standing by my side and the individuals who have encouraged and supported me through this entire period.

Dedicated to my exceptional parents: ***Shahid Rasheed & Razia Shahid****, my beloved sisters and friends whose tremendous support and cooperation led me to this accomplishment*

Abstract

The destiny of social media images depends upon their popularity: some of the uploaded images/videos get a lot of fame among people while others just get completely unnoticed. Why is this so? This work addresses this question, discusses all the features related to social content that are responsible for its popularity or negligence and also propose a system to predict the popularity of the content for the span of 30 days before actually uploading the content on any social media platform. There are some common features in the social content that gets fame, in this research work we have evaluated the effect of different features on the popularity score of the content. The proposed model predicts the popularity score in the form of number of views for the next 30 days after uploading the content. The content popularity score can be used by companies to improve their marketing strategies, targeting the right audience sagaciously, managing the resources efficiently and making the strategical decisions. In research work, the detailed methodology is discussed to design a model that can perform the task of Image Popularity Prediction (IPP) efficiently. A critical analysis is also performed on the results obtained from single features, combinational features and features obtained by applying different techniques. The best results are achieved by using Linear Discriminant Analysis (LDA) technique, which converts the higher dimensional features into six dimensions. This provides 8.03 value for tRMSE and spearman correlation of 0.76. This research work manifest that the features related to the image context i.e. user features and photo features etc. outperforms other features related to the content of the image.

Key Words: *Image Popularity Prediction, Prediction features, prediction Techniques, Social Content popularity, IPP*

Table of Contents

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

Social media and social forums are a desideratum in everyone's life. Day seems to be incomplete without using social platforms like Facebook, Instagram, Twitter etc. 60% of the world population get the internet access that is why every business and organization is using this platform to reach maximum clients and gain the benefits as much as possible. Number of entities on social media are increasing immensely day by day. Money making is very easy and simple through these platforms. Social platforms allow its users to interact, upload and share content. People upload videos and photos on many platforms and this number is increasing astronomically. Pictures are uploaded from all around the world because everyone has access to cameras through smartphones and other devices. It is easy for anyone to capture any incident or scenery immediately.

Some photos or videos are liked by millions of people and others do not get fame among people, all of this is because of some psychological factors present in human beings. If we get to learn the factors which are effecting the popularity of a photo or video among people then this data can be used to get more and more popular. This data can be used by the companies to understand the behavior of the customers on their products so, that they can target their audience sagaciously, manage their resources efficiently and make strategical decisions.

The vastness and power of social platforms attract ample amount of researchers from every domain to explore the phenomenon of information diffusion. The information is shared on social platforms in the form of images or videos and then propagated among people, this is known as information diffusion. Some information became popular and get noticed widely whereas some information is disregarded due to the unpopularity among people. Information diffusion phenomenon introduces a lot of research topics one of which is Image Popularity Prediction (IPP).

Social media analysis is done by using IPP. It can be used in any application which requires to predict the popularity of an image that is to be shared beforehand. The prediction of engagement level for an image can be used by companies to improve their marketing strategies. The ability to predict the popularity of images can surely help in several important applications such as the applications which requires content annotation and retrieval, it can also be used in the domains like content distribution and for advertisement purpose [1]. There are many other applications where image popularity data can be used like in social media marketing, brand and product marketing, to predict the popularity of political parties among people etc. The popularity of an image is usually defined on the basis of its likes, comments, shares, views, clicks etc. Understanding the psychology of people is a crucial task, what they like and what they don't like, what factors make an image/video famous among people and what factors are responsible for the ignorance of an image/video. Therefore, we need to work sagaciously to predict the popularity of an image and the most effective way is to understand your audience.

When an image or video is uploaded on any social media platforms there are some hashtags given to that image/video, which helps to place the image/video to a category according to the hashtag. These categories makes the data more discoverable, better chances of popularity and targeting the right audience. A lot of research work has been done in the past on the prediction of images popularity. Researchers have tried different techniques and approaches to predict the popularity in their own way. There are different data sets of images and the related information on which useful work has been done by many people. People performed this task by extracting different features such as sentiment features, visual features, object features, context features, user features and many other features. From the previous work it is understood that the features and statistical metadata is responsible for the positive and negative impact of an image/video on the minds of people. Previous studies predicting the popularity level lack considering the fact that the popularity will change over time. Figure 1.1 shows different social platforms.

Fig 1.1: Different social platforms [39]

1.1 Motivation

The destiny of social media images depends upon their popularity; some of the uploaded images/videos get a lot of fame among people while others just get completely unnoticed. Why is this so? The motivation behind my work is to address this question and discuss all the features related to an image/video that are responsible for its popularity among people. To run a business online the most important thing for your business is the marketing. With the right content you can connect with your customers, increase awareness about your brand, and boost your leads and sales. There must be some common features in the images/videos that gets fame, if we get to learn the pattern or features that are responsible for the fame then we can predict if media will get fame or not before actually uploading the media on social media platform. With the right content and features we can create a very strong marketing strategy.

1.2 Problem Statement

There is a need to combine prevalent methodologies, machine learning techniques and the types of features that are used by researchers in the past to predict the popularity score. In this way we can design a more efficient model to predict the pattern of popularity of an image over time. Most of the work is done to predict a single popularity score for social content but there is a need to predict the popularity score over time for any content.

1.3 Aims and Objectives

Major objectives of the research are as follow:

- Our aim is to predict popularity score for a specific time span rather than predicting a single popularity score.
- This work combines different techniques used in past to address the challenging task of IPP over time.
- It assists applications related to advertisement campaigns, recommendation systems, social media marketing and many more.

1.4 Structure of Thesis

This work is structured as follows:

Chapter 2 describes different features extracted by researchers in past for IPP.
Chapter 3 gives review of the literature and the significant work done by researchers in past for the task of image popularity prediction. It also explains the features extracted by researchers and the datasets that can be used.
Chapter 4 consists of the proposed methodology in detail. It includes three main modules: sequence shape prediction, scale estimation and the used algorithms.
Chapter 5 All the experimental results are discussed in detail with all desired figures, tables and performance measures.
Chapter 6 concludes the thesis and reveals future scope of this research

CHAPTER 2: IMAGE POPULARITY PREDICTION (IPP)

Two most important novel inventions of the last decade of 1880's are telephone and radio. These technologies still in use but the new technologies are more sophisticated than the previous ones. In 20th century technology started changing rapidly. After the invention of computers in 1940s researchers started working on the connection and communication between computers which leads them toward networking. This thing later leads to the birth of the most amazing thing called Internet.

In the earliest forms of internet emails were also developed to help people communicate. Home computers started to become common in 1980s and the social media became more sophisticated by this time, Internet Relay Chats IRCs were famous this 1990s. In 1997, Six Degrees was the first recognized social media site that allows people to make other users friends, upload profile picture and in 1999 first blogging website came into being and its still a popular sensation. The invention of blogging open new doors for networking, social media industry start to explode in popularity among people. In early 2000s websites like LinkedIn and MySpace gained prominence. In this same time Flickr Photobucket started facilitating people with online photo sharing facility. In 2005 YouTube came out with an entirely new way of communicating and sharing content among people despite of their distance. By 2006, most popular social networking platforms came into spotlight, Facebook and twitter both became available to every person of the world in a very small time. These platforms are still very famous among people. Today, we have enormous variety of social media platforms, most of them can be linked to share content of one platform to another platform which means there are very less restrictions. This allows maximum number of people to connect to each other without sacrificing the perks of one to one communication between two people. We cannot imagine the social networking platform in the next 100 years or next decade, because the way things are revolutionizing in social networking it can be speculated that social platforms will exist in some form as long as the human beings are living.

It has been a concern of humans from centuries to communicate and interact with family and friends because human beings are social animal they cannot live without interacting with other people, communication make their relationships more stronger. When face to face discussions with others became impossible or inconvenient people come up with different innovative solutions to interact.

2.1. Social Platforms

Social media can be defined as a new way of electronic communication through which people can share ideas, information, personal messages and videos etc.

Fig 2.1: Social Media [40]

How has the social media effected the lives of billions of people? Social media has effected the lives of people immensely, social platforms have become virtual gathering place. Now a day's people prefer not to go out and connect up with people on social media platforms. Social platforms offer people to have online communication through email, real-time chatting, blogs etc. Here a review is given for the most popular social media platforms of 2021:

2.1.1 Facebook:

Facebook was launched in 2004 by Mark Zuckerberg a Harvard student, it has about 1.7 billion users from all around the world which includes 69% of adults from US.

2.1.2 Reddit:

Reddit was launched in 2005 for the purpose of news sharing. Reddit provides a platform that is combination of social commentary and news aggregation. It has 300 million users on its platforms, who up-vote or down-vote a post.

2.1.3. Twitter:

Jack Dorsey founded twitter in 2006. It is a micro blogging website, around 22% US adults are on twitter.

2.1.4 Instagram

A Stanford graduate Kevin Systrom founded Instagram in 2010 as a pho-sharing platform. Later on Facebook owner purchased snapchat in 2012, it has about 1 billion users on its platform. About 75% of US adults from the age of 18 to 24 used Instagram in their daily routine.

2.1.5. Pinterest:

IPhone app developer Ben Silbermann founded Pinterest in 2010 as a visual pin board. In 2009 it became a public trade company and it has 335 monthly active users.

2.1.6 Snapchat:

A Stanford student Trio founded snapchat in 2011. It was the first video sharing platform that introduced the concept of stories, short videos in series, filters, digital effects and connect up according to location. About 73% of US adults are on snapchat.

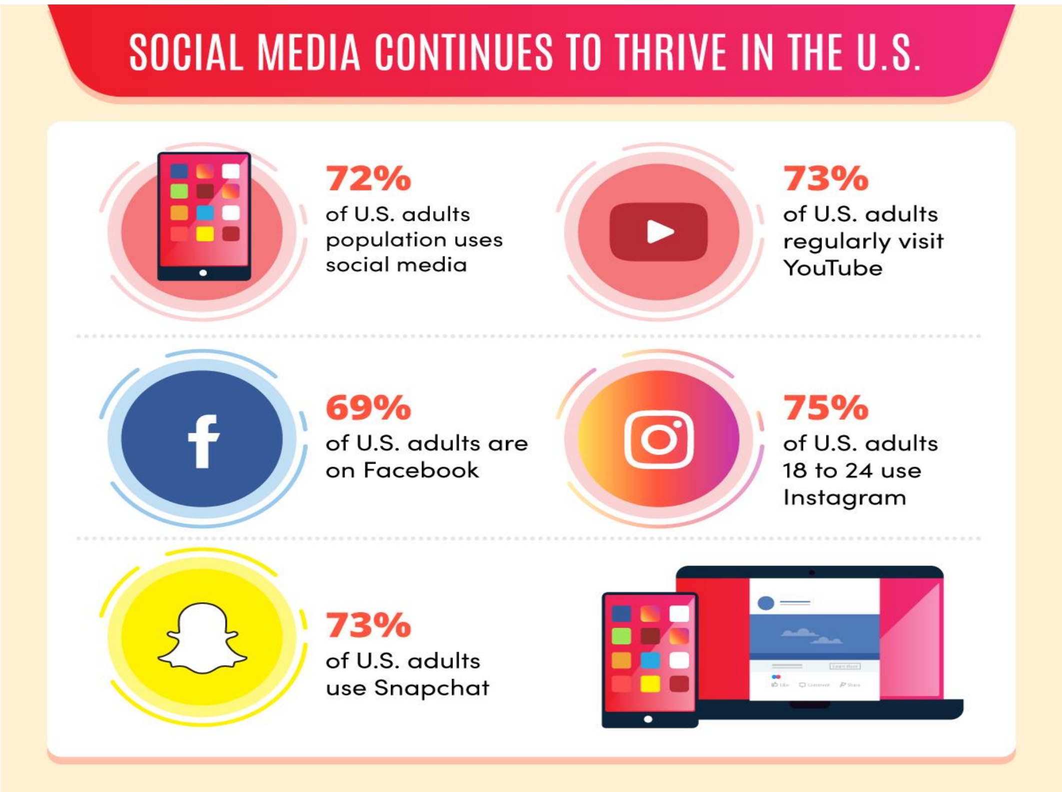

Fig 2.2: Usage of Social Media in US [40]

2.2 Businesses on Social Media

Social media began as a laptop or desktop experience that expanded to mobile phone, tablets as cellular services. Cell phone capabilities expanded this to smart phones and high speed internet connections became available to everyone in home, office, public places etc. These use everyone take their community with themselves where ever they go in the form of social media apps that run on the phone. Business took advantage of this advancement of technology of customer mobility.

Initially, social media platforms aim to help people connect with their family, friends, colleagues and like-minded people they might never meet in person. Technological improvement especially in smartphones leading to in-phone cameras, shifted the focus of app developers to videos and images. In addition to real time messages now users can have real time virtual experiences. In this era, it is very to make customer base from all around the world by using social media platforms. The companies easily make their customer base of millions of people by using business applications of Facebook, Instagram, twitter etc. Companies get access to trackable user data on the social platforms. These days users don't just log into the social platforms to browse, they give their personal information to the platform related to their name, location, likes or dislikes, their community, people they know, in short they provide a vivid picture to the marketers about them with the help of which they know how to target this customer. In early 2006, Facebook started placing ads on its platform, ads were enabled by twitter in 2010, and other apps like TikTok, Pinterest, LinkedIn, Instagram etc. have also started to monetize their services by sponsored ads.

In addition to placing sponsored ads companies came to realize the potential social media presence and activeness. Social media marketing is paid whereas sharing engaging content, information or entertaining the audience on social media platforms is another way to organically increase your customers without paying for it directly. Companies used organic social media marketing strategies to increase their customer base, target the right audience sagaciously, increase brand awareness, increase conversions, generate leads, learn from competitors and connect up with their customers. Due to the social media the reach to customer increases immensely and marketing professional work on this vertical. As marketers are no longer bounded to the old means of marketing due to social media and in this way social media marketing came into existence.

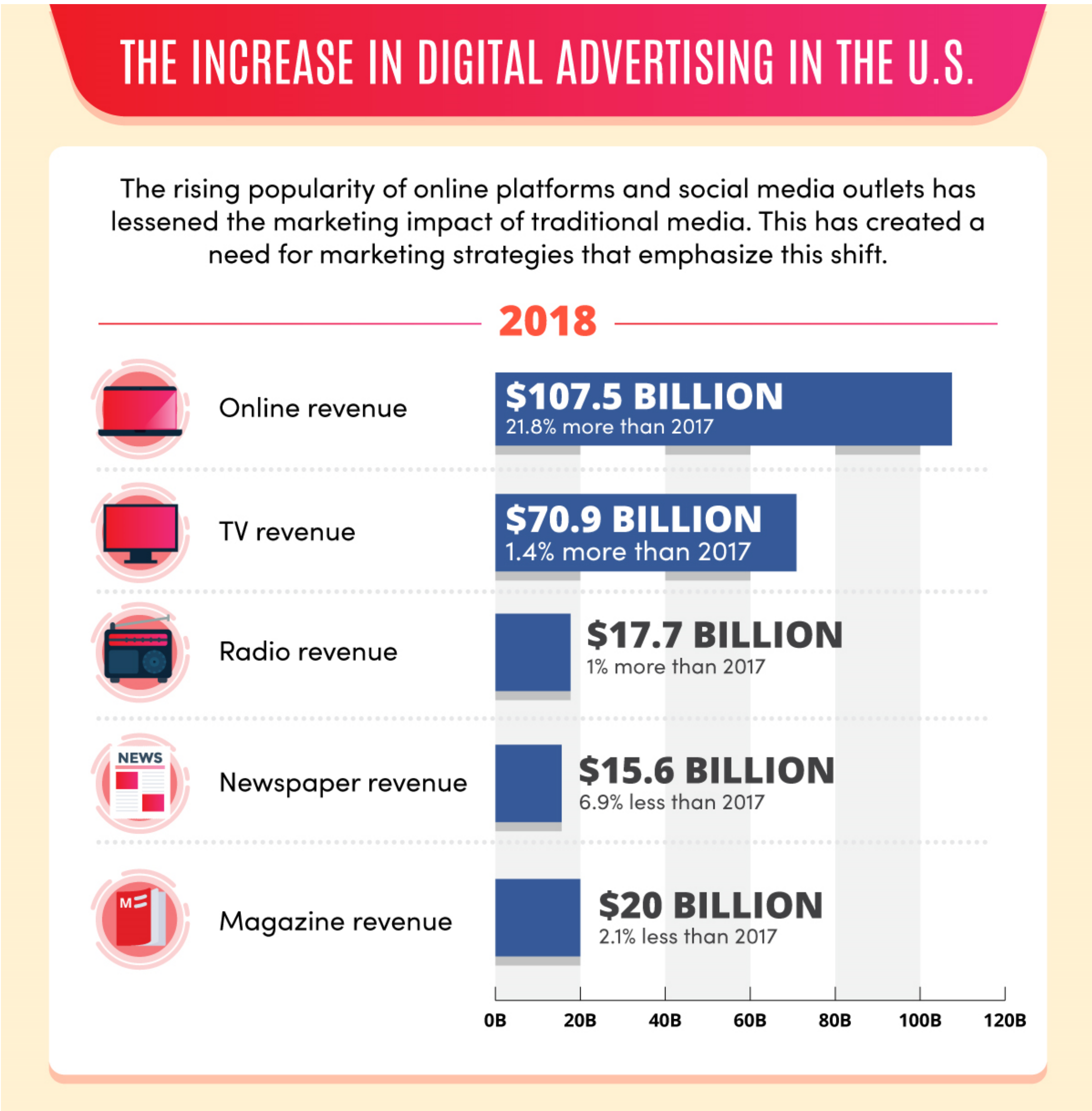

Fig 2.3: Most Revenue generating Industry [40]

Best way to take most advantage of social media is to leverage your audience. For this there are different strategies one of which is to engage influencers to share their product reviews and offers with their followers. Social media marketing is done by using endorsements and reviews by influencers who are viewed as the experts in their niche. Social platforms allow social marketers to target the right audience by placing ads in demographic groups. This can be useful in making brand awareness among the potential customers and generating leads for a specific product.

2.3. Importance of Engagement in Businesses

Social media engagement of a post can be defined in various ways by using views, comments, shares, follows etc. Social media platforms provide the services of analytics with of help of which all of these engagement parameters can be measures. The engagement analytics provides the opportunity to marketers to target the right audience or group of customers. User habits can be used for making long-term marketing strategies. For example the information related to time when maximum users are activated can be used to find the best time to post content to give it best chance of getting views. Another way used by marketers is to see cultural trends, for example if there is a cultural or religious festival in a country people will buy formal wears rather than casual wears.

2.4 Need of Engagement Prediction

There is a need for marketers to hunch about the popularity of their product or brand and how the marketing impact their customer. With the help of social platforms and tools we can have access to the data that can predict the engagement score for a social content before even actually uploading it on any social platforms, this can help marketers to make more effective marketing strategies. By knowing the engagement score marketers can test their hypothesis and find the right opportunities to engage maximum audience. With the help of engagement score, marketers can turn their hypothesis into prediction and prediction into effective social media marketing. It means it is very important to learn how engaging or image or video is for social media marketing. Image/video selection and description selection process must be very sophisticated. But how could a person know which image or video will get fame among people? We can predict which image or video will get fame by leveraging the power of Artificial Intelligence (AI). Businesses that are building their social media presence can take advantage of engagement prediction tools to engage and entertain their audience and making strategical marketing decisions.

There are some tools designed for marketing purpose that predicts the engagement score for an image or video before uploading the content on any social platform. Some of them are briefly explained below:

2.4.1. Post Intelligence (PI):

It is a rebranded startup that was launched in 2017 known as Pi, it is an AI-powered prediction tool for marketing. It can predict the engagement score the even boost the engagement score by writing the post. Pi works by using social media trends and user data related to post to predict the popularity of the image in the form of likes, views and comments etc. The engagement by predicted by a single number ranging from 1 to 10. 1 means the lowest engagement and 10 means the highest engagement.

2.4.2. LikelyAI

LikelyAI predicts the popularity of images on Instagram by using Artificial intelligence. The prediction is done by using data points and patterns like color, objects, shapes, emotions etc. The algorithm deeply evaluates every image by using AI to find the best one that can get maximum likes, views or comments.

2.4.3. MIT Prediction System

The researchers at MIT designed an algorithm in collaboration with eBay lab and DigitalGlobe to predict the popularity score for an image. It predict the popularity on the scale of 0 to 10. They designed the system by using 2.3 billion Flickr images, the algorithm provides the result by taking into account features like color, texture, user data, social context and content etc.

CHAPTER 3: LITERATURE REVIEW

This research work consist of the literature work from books, journal articles, conference papers and case studies. The main objective of this section is to collect, synthesize and organize the existing knowledge related to image popularity prediction on social platforms. I have reviewed several papers in which work has been done to predict the popularity of content before actually uploading it on any social platform. I have searched different databases ResearchGate, Springer Link, Association for Computing Machinery (ACM), IEEE Xplore, and some other journals for computer science. For the purpose of searching the relevant articles the keywords used are social content popularity prediction, image popularity prediction, prediction features and popularity prediction techniques. The reference list of each article is checked for any relevant potential research on this topic. The span of publication taken into account is from 2001 to 2019. The PDF and full-text documents are used that contains complete data about the keywords and literature. A substantial amount of papers have been considered to contribute in this study.

In the recent years, the topic of image popularity prediction has attracted ample amount of researchers towards it. Image diffusion has become a salient topic among researchers from all around the world. Significant amount of effort is spent by a lot of people in searching new and better ways to predict the image popularity among people. Different works agree on different features to provide best predicting results. In this section we will briefly discuss the work contribution by different researchers by utilizing different features.

A lot of work for image prediction is done by using textual data of the image but some researchers have worked on the task by using digital image processing techniques and classification methods on the image context, image content and features. Different strategies have been used by different researchers. It can be perceived from the previous work that the social features of an image are more powerful than the visual features of the image. Best features for an image prediction actually depends upon the case, if a user has just joined the social media than there is no metadata for that user like the descriptions or tags etc. In this case visual features would be a better approach for image prediction.

For image prediction a popularity score is computed for every entity of data set which varies in the definition such as image views, likes, comments, shares, mean views in a specified period etc. Despite different definitions every one of them have same basic pipeline structure. They

Extract features related to content or context and then calculate popularity score by using regressor. There are some companies and teams who have developed algorithms and applications that can be used to check the popularity score of your image before uploading it on any social platform. An algorithm is designed by the collaboration of the scientists from the lab of ebay and the scientists from the artificial intelligence lab of MIT, which predicts the popularity score in between 1 and 10. This algorithm uses all features of an image related to color, objects, textures etc to predict the popularity score. The prediction not only ends on the popularity score people have worked far more than that. An algorithm is developed by a company named beautiful destinations which will predict the number of likes and comments before uploading the image and it will even predict the type of comments your image will receive. This algorithm works by making correlation of image data gathered from different platforms. This algorithm is capable of analyzing the effect of editing on a picture and the audience reaction to those changes.

Totti et al. [3], worked on the aesthetic features of an image. Aesthetic features are the features such as blurness, statistics of channel colors and aspect ratio. The 85 object classifier is used for these features together with the content features. It can be seen that image context and user features produces very promising results. Khosla et al. [4], worked on both simple and low-level visual features such as color, contrast, GIST, color patches, LBP, texture, gradient features etc. The work also consider high level features like determining the objects present in the picture and network activation to predict the numbers of views and popularity an image will get by implementing the state of art technique of CNN or LinearSVR for classification. In [16] researchers have used DNN model on the data set collected from an Instagram page of Indian lifestyle magazine to predict the popularity of posts to be done in future. To learn the weights in the DNN model they have used mini-batch-gradient-descent method, cross-entropy is used as an objective function. Their model achieved the accuracy of about 70%.

McParlane et al. [2] has proposed an algorithm that uses user, content and context features of an image to predict the number of comments and likes on a data set MIR-Flickr1M. Their model do not predict on the basis of high-level features of the image. They used SVM for the classification using RBF kernel. In [5] Gelli et al. has proposed an algorithm to use the combination of visual sentiment features, context features, content features and user features to predict the popularity of images on Flickr. They did it by doing classification using Convolutional Neural Networks (CNN) and Support Vector Machine (SVM). In [6] researchers has worked on the three important factors for predicting the popularity on any social platform: social context, visual content and textual information. They have used the Flickr data of 1.5 million images that is uploaded from Application Programming Interface (API). Their popularity score depends on the number of views an image get. They perform the classification by using the Ranking SVM classification methods and check its implementation under different circumstances. Keneshloo et al. [9], proposed an algorithm to predict the number of views a new article will receive after the publication in the first day. They used the news articles from The Washington Post and extract the temporal and contextual features along with social and metadata features of the news articles to predict the popularity.

In [7] Hu et al. have worked on Yahoo Flickr Creative Commons 100M (YFCC100M) by implementing Caffe deep neural network framework [8]. They have extracted different visual features for every entity of data set. Then used the extracted features for several learning approaches like unary and multimodal approaches to predict the popularity in the terms of number of likes on an image. I their experiment they choose 10,000 images from their data set and uses their tagline information in addition with the visual features of the images. After the experiment they concluded that the tag feature outperforms other features. Trancinski et al. [8] have predicted the popularity of a video by proposing a regression method which uses the number of view of the videos to train. The model uses Gaussian Radial Basis Functions with the Support Vector Regression method. They have used 2400 videos from Facebook and YouTube as a data set, for which they analyzed the visual and social features. They concluded that the social features turn out to be stronger than the visual features for the popularity prediction of a video.

Mazloom et al. have proposed an approach of predicting the image prediction by using the preferences of an individual user regarding different items. They perform their experiment on 600K posts in that are made on Instagram which are related to the tourist places in the Netherlands. Their purpose was to predict popularity shared by a user related to an item by using the textual and visual context of the image [10]. They also have has proposed an algorithm for popularity prediction which is category specific. They have used the textual and visual content like scene, action animal and presence of people etc. They performed this experiment on a data set of posts taken from Instagram which were about 65K. Their model works on low-level features, conceptual features, visual sentiment features, bag of word features, word to wee features and textual features [14].

Wu et al. [11], have presented a novel framework for prediction known as Deep temporal Context Network (DTCN). This proposed system is for the sequential popularity prediction. They perform the experiment on 680K images that were uploaded 3 years ago on Flickr, the data set used was TPIC17. Their work concludes that the DTCN approach outperforms all other approaches used on the data set previously. In [12] Fernandes et al. have spent two years for collecting data from the Mashable website of 39000 articles. They extract total of 49 useful features that can predict the popularity an article will get in the form of number of shares. Five classification methods were used by them which are Adaptive Boosting, Random Forest, SVM with RBF kernel, Naive Bayes and K-nearest neighbour method. They receive about 60% of the accuracy. Nwana et al. [13] proposes an algorithm for the popularity prediction of the youtube videos in a campus.They uses two approaches for this purpose: Social approach and consensus approach. They uses simple caching framework to measure the performance of their system. Jheng et al. [15], have developed a drift based predictor of popularity. They have combined multiple classifiers to train their model. They have worked to predict the popularity of multimedia on social platforms in a micro blog. In [17] researchers have designed a real-time popularity predictor that is based on the user feedback. Their developed algorithm uses the user's feedback in accordance to time. The data set is collected from the Instagram of 500 bloggers, in total 100,000 posts and the updates of about 1, 00,000,000.

The research work done the literature can classified into the categories based on the data type they are working on:

3.1. Visual-Sentiments-Based:

A lot of work has been done for the analysis of multimedia sentiments of the images on social platforms. Emotion Wheel of Plutchik provides 24 basic emotions in an image [18]. A SentiBank named ontology of visual sentiments have been presented by Broth et al. [19], they have trained ANPs with 3,244 detectors by combining the local and global features of an image. Convolutional Neural Network has done a breakthrough in the analysis of visual sentiments. Chen et al. [20] have worked on the model of Broth et al. [19], by replacing the SVM model with the CNN model to achieve better accuracy. Localization of an object and sentiment classification can also be done a hierarchical model proposed by Chen et al. in [21].

3.2. Text-Based:

Yu et al. [22], have used the text, user and temporal data of tweets from twitter, to check the number of retweets. Hong et al. [23] have used the tween content related to the tweet's topical information like temporal and user features. Zaman et al. [24] have used the tweet content and user information to predict the number of retweets a tweet will get, by using the user patterns.

3.3. Multimedia-Based:

McParlane et al. [2] have predicted the popularity using user information, image content and context. Khosla et al. [4] have worked by using image content and social context. In [7], [25] the authors have used the user information, hashtag and low-level features to predict the popularity. [27], [28] have used the images content and social content for the prediction. In [3] the authors have used aesthetic and semantic features of the images taken from Pinterest. In [26] Foilet et al. have trained the model by using image and user information. Fioletet al. have classified the prediction of popularity by using the network based features in [29].

3.4. Type of Problem:

The problem of popularity prediction can be categorized into three main categories: Regression, Retrieval and Classification.

- **Regression:** The popularity of a post is quantified.
- **Retrieval:** Ranking of images between 1 to 10.
- **Classification:** Popularity is classified in the form of classes.

The work done in [22], [25] can be categorized as Regression one. The papers [2], [3], [22], [24] have formalized the problem of popularity prediction into the category of classification. Authors of [4], [5], [26], [27], [28], [29] have solved this problem as a retrieval one.

Researchers in past have used different approaches to achieve the task of popularity prediction of image/video. Some of the famous prevalent approaches taken by the researchers in past are discussed here. In [5], researchers have achieved the task of popularity prediction by using a different approach shown in Figure 3.1. They have extracted four types of features which are user features, context features, object features and sentiment features for the images in the data set. All of the extracted features are then used by their model to predict the popularity. Then a score is assigned to the images for example 50 views/day, that score predicts the popularity of an image before uploading the image/video. They have trained their model with two different scenarios. In the first scenario, the images used as data set are 25K but one picture is taken from each user of Flickr. In the other scenario, they have used all the images in the data set that are taken from 25 users of Flickr, 10K images are selected for each user. In this way a specific user can predict beforehand that what type of picture uploaded on this profile will get more fame. Their experiments conclude that the popularity of an image has correlation with the sentiment features but user features still provides better results for the prediction. Combination of these features provides better accuracy. In case where user data is unavailable sentiment features can play significant role in the prediction.

In [33] Almgren et al, have proposed a three step approach for the prediction of popularity score shown in Figure 3.3. They have studied the effects of image semantic on the popularity of the images. They have used the clustering technique and natural language processing to extract the semantic of an image from the caption given to that image. In the first step of this approach, pre-processing is performed on the captions to find out the keywords that can explain the entire caption and categorize it. In the second step of this approach the extracted keywords in step 1 are converted in numerical form by using Word2vec. This is done to check the keywords and similarities between them. In the final step 3, a keyword vector is generated by using the K-means algorithm. Multiple semantic of an image are represented by this vector.

Zohourian wt al in [32] have proposed their own methodology for popularity prediction shown in Figure 3.4. First step of their methodology is the aggregation of data from social platform. Second step is the extraction of features. They have extracted the common, time, visual, textual, figures and video exclusive features from the images that are collected as the data set. As the third step some pre-processing is done on the extracted features like catering the missing values and nominal to numerical conversions etc. Next two steps in the schema is for the prediction purpose for the test data set. In step 4 different regression methods are used to predict the popularity score. They have used Linear, Local Polynomial. SVM Linear and SVM Kernel regression methods. The final step is step 5, in which the classification methods are applied on the features to predict the class of the popularity for the image/video. Classification methods used by them are ID3, KNN, Naive Bayes, Random Forest and AdaBoost. Their experimental results conclude that the decision tree algorithms and local polynomial regression methods outperforms other methods in the achieved accuracy.

In [34] Meghawat et al. have proposed a methodology for multi-modal social media popularity prediction shown in Figure 3.2. The output value of this model is the predicted popularity value. This value is predicted by using both image content and context information. They have extracted total 15 features of an image. Fourteen features are extracted from CON-SOC information that includes the information of picture ID, user ID, postdate, comment count, has people, title length, description length, tag count, average view, group count, average member count, title, tags and description. One value is taken from the prediction output which in total became 15 features. These 15 features are passed to a Convolutional neural network. The model is made up of two dense layers and two convolutional layers. The output of this model is final predicted value for the image/video.

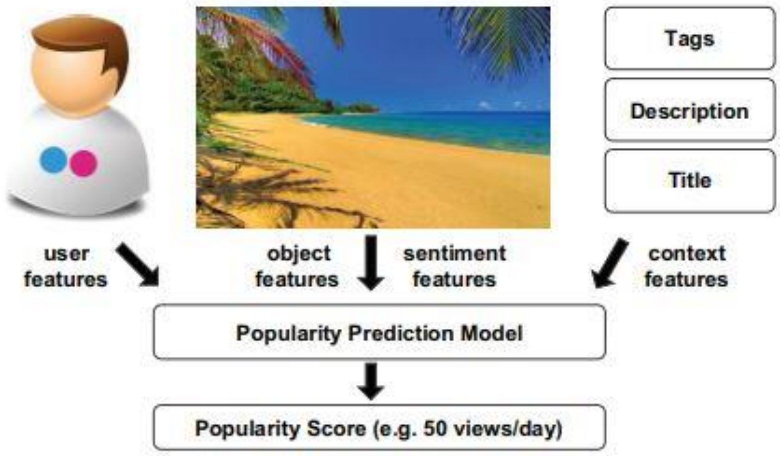

Fig 3.1: Schema of approach used by [5]

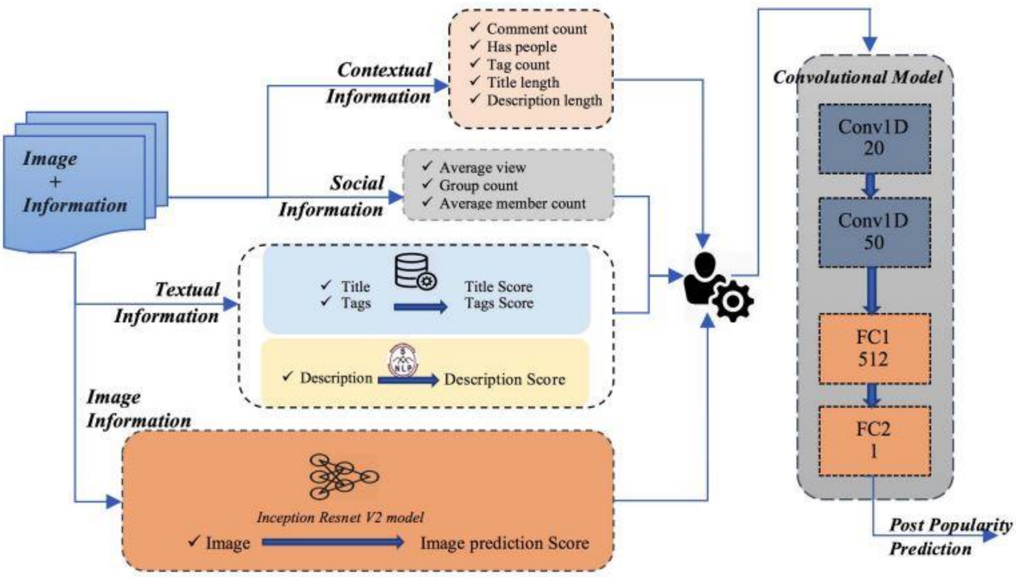

Fig 3.2: Schema of approach used by [34]

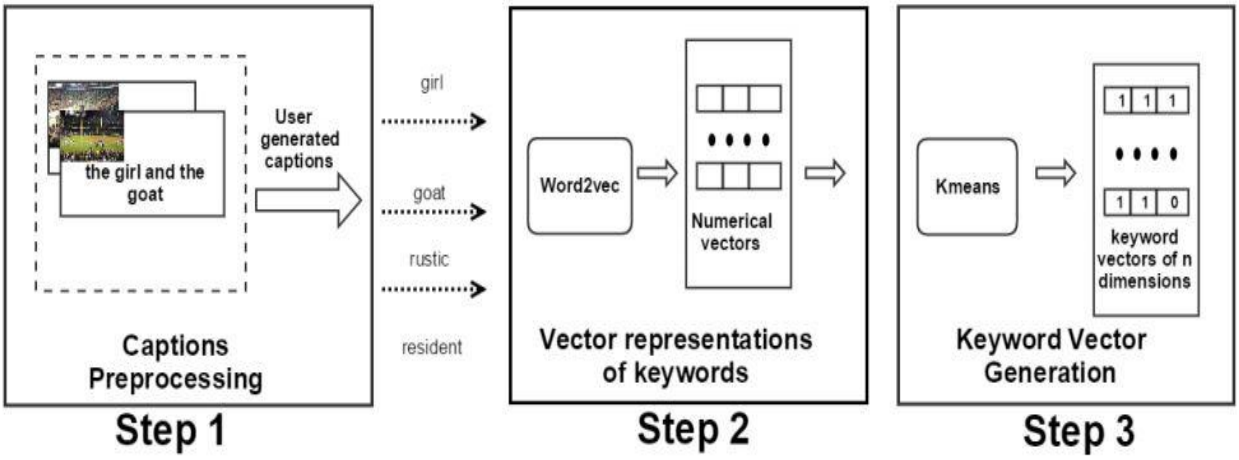

Fig 3.3: Schema of approach used by [33]

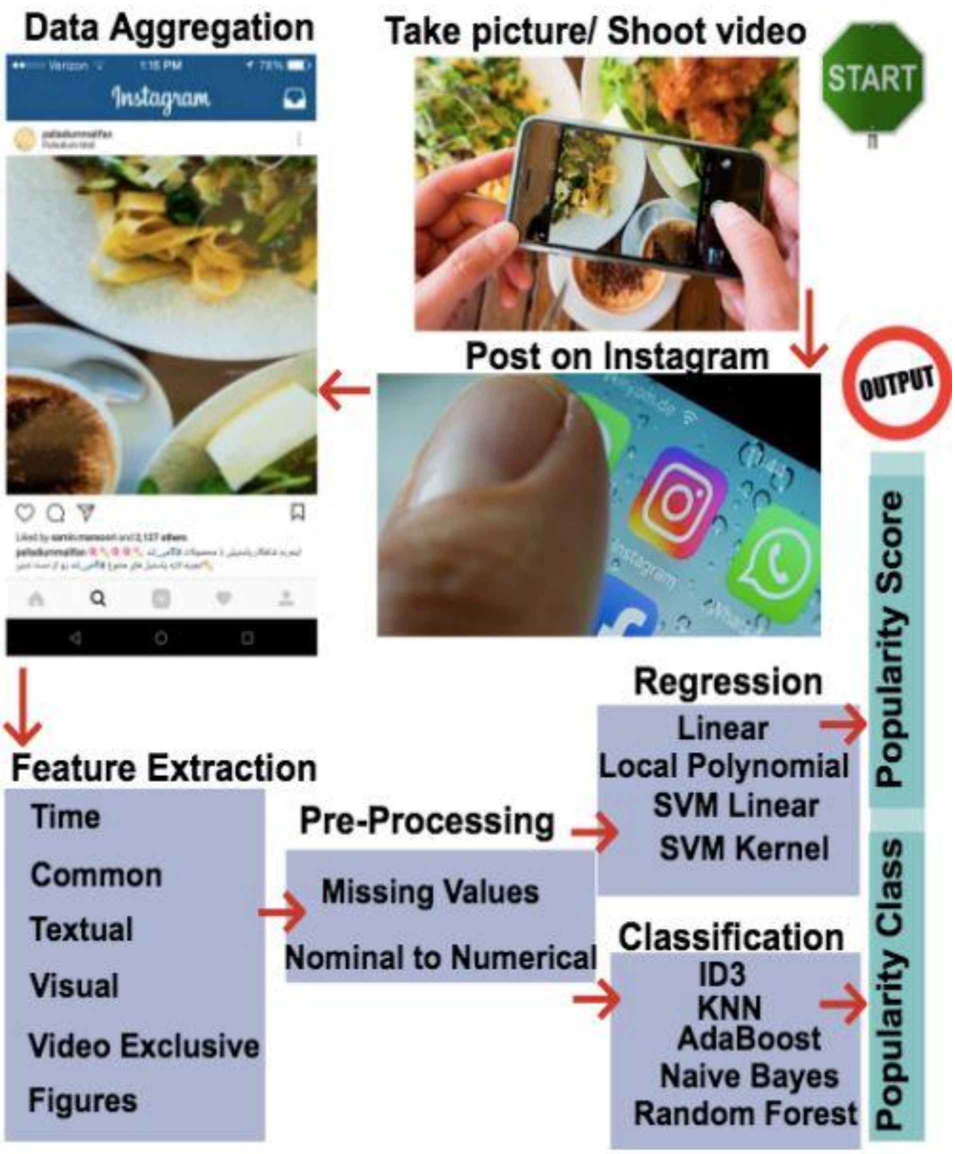

Fig 3.4: Schema of approach used by [32]

3.5. Features Extraction

There are ample amount of features that can be extracted from the image content and context for the popularity prediction. Each type of feature has its own effect on the prediction process. Some of the features that can extracted for popularity prediction are:

3.5.1. Time-Features:

In the extraction of time features we consider all time faces related to the uploading of an image or video. These timely features can include the information regarding season and month in which the image/video was uploaded. It can also contain the information that the image/video was uploaded at which time of the day or which day of the week, was it a weekend or any specific event holiday.

3.5.2. Object-Features:

Researchers in past have proved that visual content of an image/video immensely effect the popularity. So we can extract the features related to the objects present in the images or video. This form of extraction can be done by using Convolutional Neural Networks (CNN), which were proposed and updated by many researchers. If we use a deep CNN that have 16 layers to extract the object features of an image, we can receive the final output with 1000 objects that are termed as ObjOut and we can receive fully connected rectified layer with 4096d representation termed as OBJFC7 according to [30].

3.5.3. Common-Features:

These are the features which are common in the images or videos. These features includes the file type of the image/video, Topic category to which the image or video relates for example it could be about lifestyle, motivation, culture etc. The height and width including the orientation (landscape, square or portrait) of the image/video is also considered in theses features. The type of feelings a person gets after seeing the post is also considered by this category. It also includes that the image/video that is posted is created by the user himself or it is copied from any other source.

3.5.4. Context-Features:

Context features of an image/video are related to the information of tags and description given for that image/video. This information can contain very useful data that can be taken as a clue to predict the number of views an image will get. The popularity of an image/video can be effected by the entities like the person who has shared it, the location or the tourist place attraction. People might be more interested in seeing any specific object, in the case the tags and descriptions given for image/video that is shared in any group become more accessible to be found by other people. The text used in tags and description is very important for the extraction of features from the context. Tags have very concise text and description is written in natural language with full explation of the entity. Here two different approaches can be used to extract context features: Extraction of features by using tags and Extraction of features using description.

3.5.5. Text-Features:

This category of feature extraction extracts all the features related to any kind of text that is given for the image/video. The features that can be extracted can be the caption that is given for the image, the hashtags for the image that categorizes the image, the text written inside the image/video that is referred as media text etc.

3.5.6. Visual-Sentiment-features:

Visual sentiment features are the emotions that a person can feel after seeing the image/video. Such features can be extracted by using the classification that is performed for visual sentiments. Borth et al. [19] has defined a Visual Sentiment Ontology (VSO) which contains 3,244 Adjective Noun Pains (ANPs). We can also use DeepSentiBank [20] that is a CNN that is pre-trained to classify images in 2,096 ANPs subsets.

3.5.7. Video-Exclusive-Features:

These are the features that are related to any video for example the music in a music that if its a global or local music or there is no music at all. Other features for this category are if there is a narrator in the video, if there is any is it male or female. Another feature can be the cover of the video which can be a logo or a frame.

3.5.8. User-Features:

It can be clearly seen by the previous work that the user features play a significant role in the prediction of the popularity of an image/video. The popularity obtained by an image/video not only depends upon the content of the image but depend on the context information. The author data is very important in the task of popularity prediction. These features can be the number followers of the user, number of groups in which the user is present, views, likes and comments. Some other user features are used by Khosla et al. in [4].

3.5.9. Visual-Features:

Visual features are the features related to the images like the contrast, brightness, median etc. These features consist of the mean values for the red, green and blue color present in the image. There are many features that can be considered from the image to predict the popularity.

Table 3.1 shows different significant research works done in past. This table evaluates the research work for the task of IPP by comparing the features used in each work. Literature work is evaluated by using nine features: time, object, context, text, social, user, visual, sentimental, low-level features.

Table 3.1: Features extracted by authors to predict popularity score

Authors	Time	Object	Context	Text	Social	User	Visual	Sentimental	Low-level
McParlane et al. [2]		✓	✓			✓			
Totti et al. [3]			✓			✓			✓
Khosla el al. [4]					✓	✓	✓		
Gelli et al. [5]			✓			✓	✓	✓	
Aloufi et al. [6]			✓	✓			✓		
Hu et al. [7]			✓			✓	✓		
Keneshloo et al. [9]	✓		✓		✓				
Mazloom et al. [10]				✓		✓	✓		
Wu et al. [11]			✓						
Fernandes et al. [12]	✓		✓	✓					
Nwana et al. [13]					✓	✓			
Mazloom et al. [14]				✓			✓	✓	✓
Trancinski et al. [18]					✓		✓	✓	
Broth et al. [19]							✓	✓	
Chen et al. [21]		✓						✓	
Yu et al. [22]	✓			✓		✓			
Hong et al. [23]	✓					✓			
Zaman et al. [24]			✓			✓			
Foilet et al. [26]		✓				✓			
Cappallo et al. [27]		✓			✓				
Yamaguchi et al. [28]			✓			✓			✓

3.6. Different Approaches

Researchers have calculated the popularity score in different ways. In [5] Gelli et al. , have calculated the popularity score (r) by using Support Vector Machine (SVR). The data set used was very large in size so, they have used Support Vector Regression (SVR) with L2 regularized L2 loss from the LIBLINEAR library package. This type of regression is used due to its scalability and efficiency with large amount of data and immense amount of entities than the kernelized version of SVR. L2 normalization is used to normalize the multi- dimensional features. The range of C for C is set in between [0.001-100]. The images are ranked in the descending order after performing the prediction. The ranked images are then compared to the true popularity score that is achieved by the ground truth value (s). The correlation between both scores is then calculated by using spearman's-rank-correlation. The range for this correlation is [-1,1]:

$$p = \frac{\sum_i (r_i - \bar{r}) - (s_i - \bar{s})}{\sqrt{\sum_i (r_i - \bar{r})^2} \sqrt{\sum_i (s_i - \bar{s})^2}} \qquad (1)$$

In [4] Khosla et al. have calculated the popularity score by using social media content. The cumulative engagement that is achieved till the download time is used as a social media content. This cumulative engagement score is normalized by using the number of days from which the content was uploaded on social platform. The score equation can written as:

$$s_i = \log\left(\frac{c_i}{T_i} + 1\right) \qquad (2)$$

(c) is the engagement score achieved by an entity (i). This engagement score can be in the form of number of comments, views, likes or shares etc. The engagement score is achieved for the entity (i) for T days since the image/video is uploaded. The popularity score (s) for entity (i) will be calculated according to the equation.2.

In [32] Zohourian et al. have predicted the popularity score by using more than one approach. They have applied four different techniques of regression on the data set to calculate the popularity score. The methods used by them are:

- Linear Regression
- Local Polynomial Regression
- Support Vector Machine
- Linear Support vector Machine

$$score = \frac{No.of.Likes}{No.of.Followers} \qquad (3)$$

They have used linear regression to tune the parameters minimum tolerance and ridge by selection from four feature including Greedy, M5 Prime, Inductive T-test and T-test. In Local Polynomial Regression ridge, numerical measurement, degree and neighborhood are the factors that are engaged in popularity score prediction. Support Vector Machine they have used kernel of type dot, polynomial, radial, anova, neural, multi quadric and gaussian combination. They have tuned the parameter C to achieve better results. In linear SVM they have used linear kernel.

In our proposed work model works on the engagement sequence s, which is given for every Flickr image according to the number of views, for the span for 30 days. In our proposed methodology the engagement sequence s is splited into two parts Sequence scale (s_{scale}) and sequence shape (s_{shape}). Sequence scale and sequence shape are defined in equation 5 and 6:

$$s = s_0 + s_1 + s_2 \ldots \ldots\ s_n \qquad (4)$$

$$s_{scale} = \max(s) = \ s_n \qquad (5)$$

$$s_{shape} \ = \left[\frac{s_0}{s_n}, \frac{s_1}{s_n}, \frac{s_3}{s_n} \ldots \ldots \ \frac{s_n}{s_n}\right] \qquad (6)$$

The maximum value in the engagement sequence s is taken as sequence scale. On the other hand sequence shape is obtained by diving the engagement sequence s with the maximum value in the sequence that is sequence scale (s_{scale}). In this way, we get normalized values for (s_{shape}) due to the division with maximum value. Therefore, these two parameters sequence scale (s_{scale}) and sequence shape (s_{shape}) can be used to obtain the original sequence s by using equation 7.

$$s = s_{scale} * s_{shape} \qquad (7)$$

In this proposed method, estimations have been performed to predict the engagement score in the form of (s_{shape}) and (s_{scale}). Equation 4 is exploited to predict original popularity. This proposed system assumes that (s_{shape}) and (s_{scale}) are independent of each other because practically it is observed that images that have similar (s_{shape}) have very different values of (s_{scale}). So it proves that (s_{shape}) and (s_{scale}) have independent relationship between them.

3.7. Data Sets

While a lot of people work on the challenge of popularity prediction of images, therefore many of them came up with private data sets. Some data sets were developed from the scratch like they have downloaded the images from a social platform at their own and some of them have make new data sets by exploiting the publicly available data sets. The publicly available data sets are developed to allow the researchers to successfully apply experimental protocols and to do some meaningful research without doing of effort of collecting data set. Data sets that are publicly available are Popularity Dynamic Data set [35], YFCC100M Data set [36], Flickr30k Data set [37], Open Images Data set [38] and there are many other private data sets made from public data sets. The key characteristics of these publicly available data sets are briefly explained in the following. Ortis et al. [31] have built and released a new data set that is based upon the images from the Flickr API. The platform of Flickr only provides a cumulative engagement score so they built a data set that provides the multiple values of daily score for the span of 30 days for each image. The data set do not involve any biasness, it also contains all social features that are related to an image such as user's information, groups, views and comments etc. This data set consist of more than 20,000 images that have been observed for 30 days after they were uploaded on Flickr. Yahoo Flickr Creative Commons 100M (YFCC100M) is a data set that contains 100M publicly available images and videos on Flickr under the license of creative commons. This data set consist of 99.2 million images and 0.8 million videos in total. It is an unsupervised data set. This data set is widely used by the research students and people who work in the domain of artificial intelligence, machine learning and computer vision. Flickr30k data set consist of 30k images that are taken from the Flickr and is publicly available on kaggle. This data set has become a benchmark for sentence based description of the images. This data set can be used for many tasks like text-image embedding, color classification, common object detection etc. many researchers have exploited this data set for their own use like extracting pictures that have one line tagline etc. Google has introduced a data set named as Open Image data set, which contains about 9M varied images. URLs of the images are provided in the data set with the labels of the images categorized into 6000 categories. This is quite a practical data set that contain real-life entities. This data set is the collaborative effort of Google, Cornell University and CMU. It contains bounding object boxes, annotations based on image level, object segmentation, localized narratives and many others.

The dataset used in this work is the one released by Ortis et al. If we check on the Flickr platform, it just provide us with a cumulative engagement score in the form of likes, views etc. So they come up with a dataset that provides multiple value by using the crawling process shown in Figure 3.5; engagement score of every day for the span of 30 days for each image. This dataset is not biased in any way, it is an extension of Social Popularity Image Dynamics Dataset. This dataset consist of approximately 30k Flickr images which are crawled for the span of 30 days from the time of uploading to get popularity score for each image. This dataset consist images and many different features related to the images. It consist of all user and photos social features that includes user information, no of user groups, mean views, comments, tags, descriptions and many others. Figure 3.6 shows some sample images from the dataset.

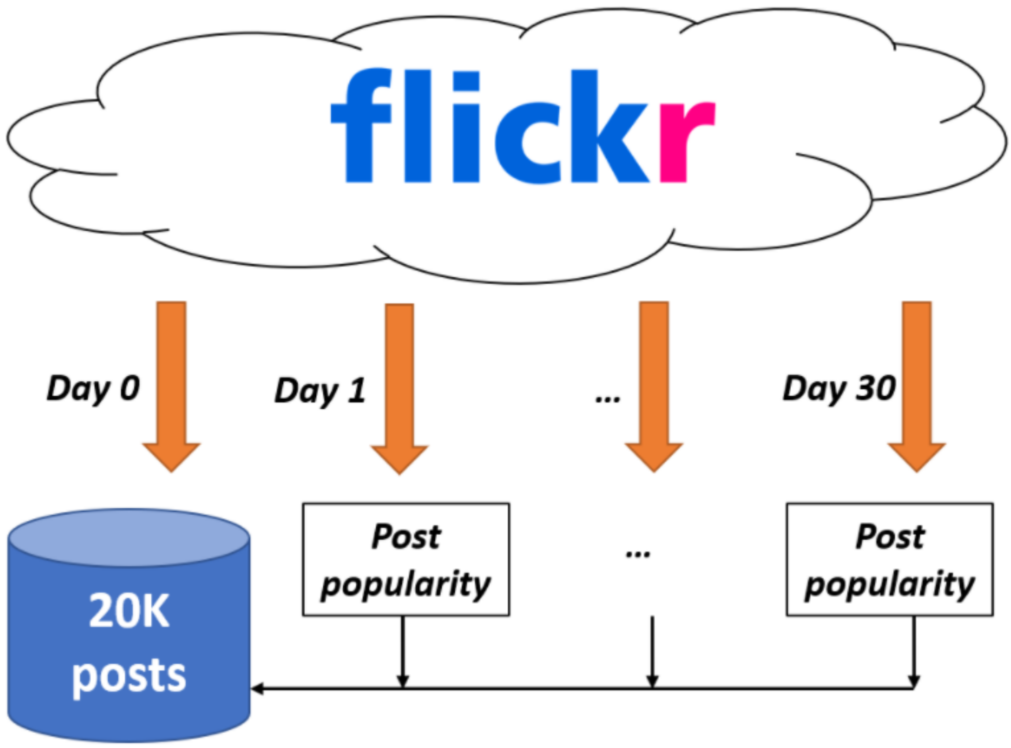

Fig 3.5: Dataset collected by Ortis et al. [31]

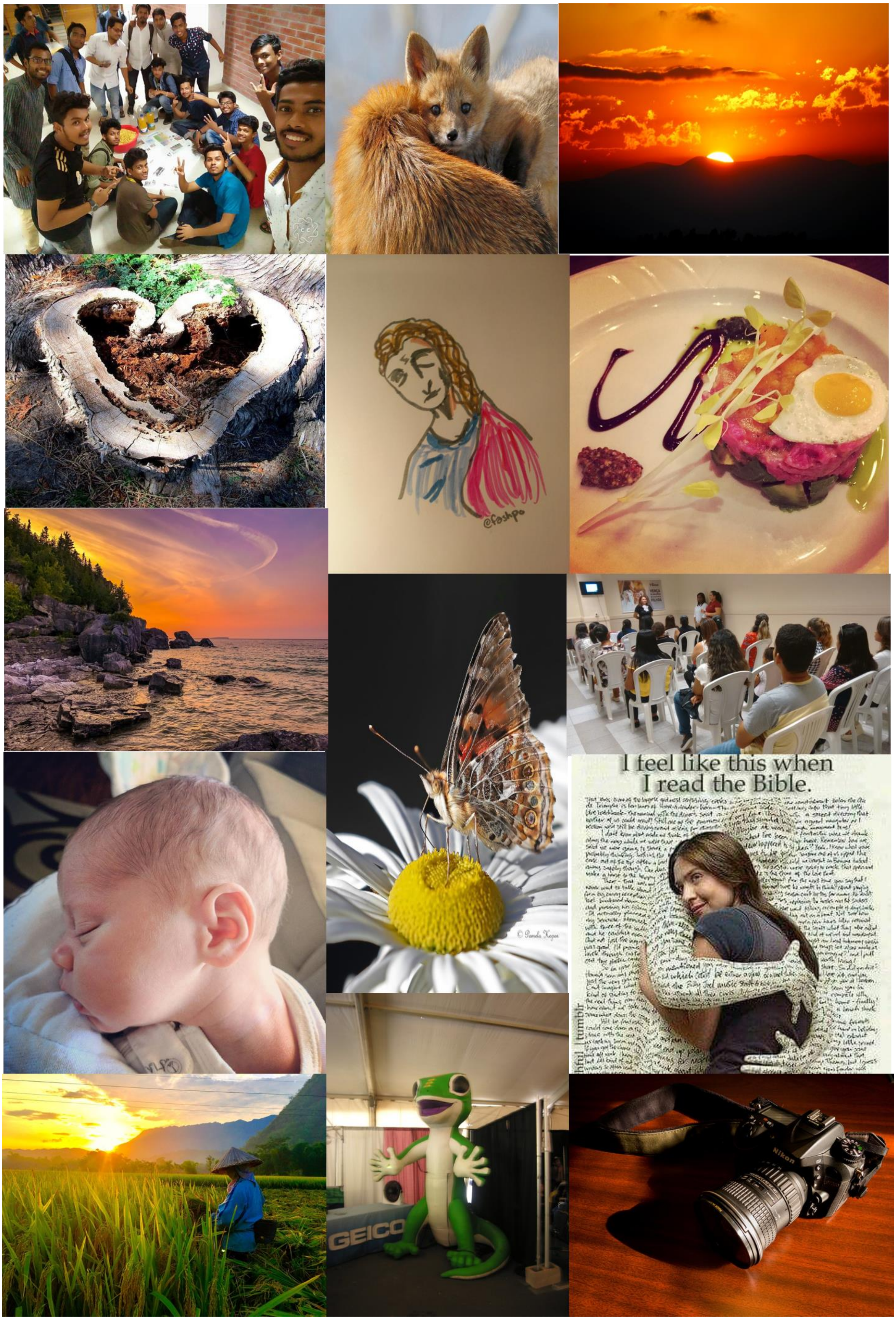

Fig 3.6: Some images from the dataset.

3.8. Research Gaps

Most of the work is done to predict a single popularity score for social content.

There is need for further investigation to address:

- Questions related to image popularity over time i.e. best features, factors that make an image popular etc.
- Some challenging problems in this domain i.e. the best techniques and models for IPP.

CHAPTER 4: METHODOLOGY

The overall schema of the proposed method is shown in Figure 4.1. The methodology is divided into two parts: Training Phase and Testing Phase. In the training phase, first of all shape clustering is performed, results of clustering gives us general shapes which are exploited to obtain prototypes. The obtained clusters are then used by classifier as the classes, to train a model to predict the cluster of the new sequence according to which the prototype will be assigned to that sequence. An SVR is trained for scale estimation, this model is trained by using a different sets of social features to predict popularity scale. In the testing phase RNDF classifier that is already trained predicts the sequence shape s'_{shape} by doing classification and assign a prototype defined by the clustering analysis. The trained SVR model gives the sequence scale value s'_{scale}. Overall engagement score for the span of 30 days is then obtained as: $' = s'_{scale} * s'_{shape}$.

The performance of the model is measured by using Root Mean Squared Error (RMSE) and Spearman's correlation. In the experiment dataset is splited into 80% training data and 20% testing data.

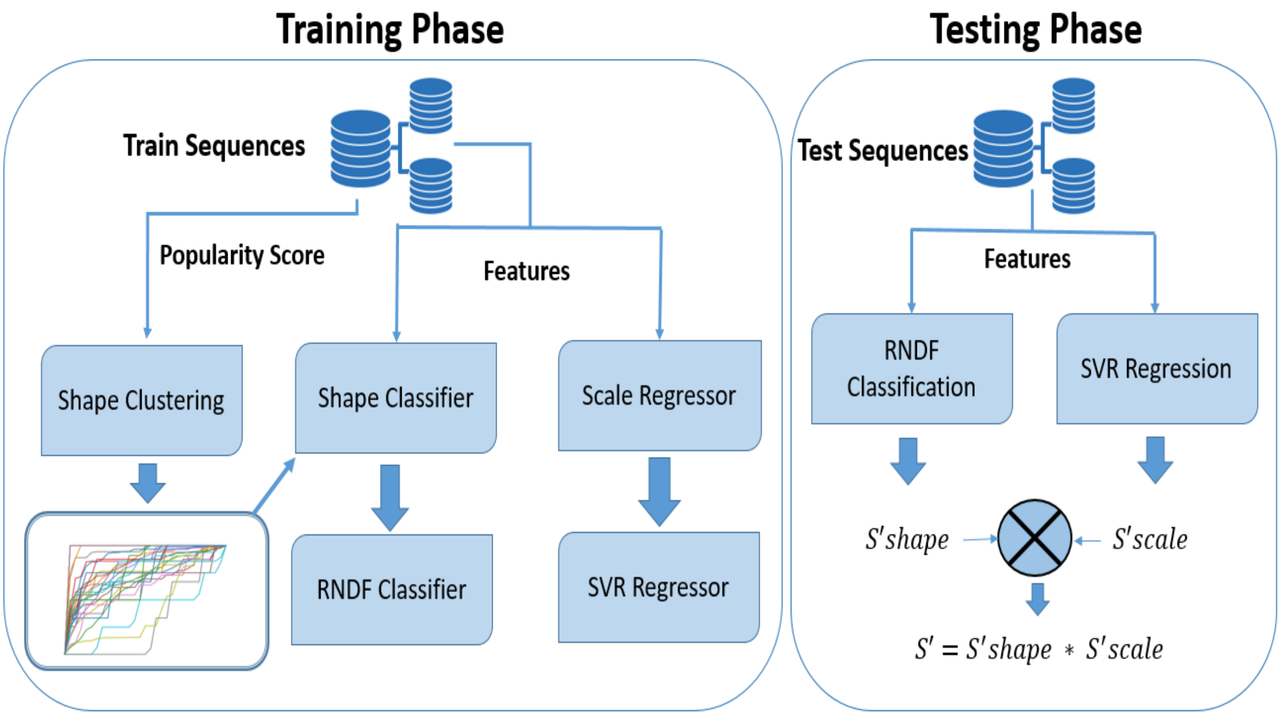

Fig. 4.1: The Proposed Methodology: Training phase and testing phase are shown separately. This model solves the problem of IPP by predicting and combining s_{shape} ands_{scale}.

4.1. Sequence Shape Prototyping and Prediction

According to the formulation of equation 3, the sequence shape is normalized by diving the individual value of popularity score with the maximum popularity score of that sequence due to which it ranges between [0,1]. This makes it easy to check for the sequence shapes that are similar to each other and group them into a single cluster. Here we have considered that all sequences with same dynamics belongs to the same group and each group has same engagement score evolution. First of all we tried to define the number of clusters in which we want to divide the dataset. We have used many different clustering algorithms like Affinity Propagation, Agglomerative Clustering, Birch, DBSCAN, Guassian Mixture, Spectral Clustering and K-means clustering to group the normalized sequences. The best results in our model for clustering are provided by K-means clustering. The resulting centroids for every cluster define the prototype of each cluster. For determining the best value for K, we have done evaluation of our model using different K values K=10, 40, 45, 50. Then the best K has been selected in our case which is K=50 for the span of 30 days. Figure 4.2 shows the centroids for each cluster, each centroid represents the corresponding prototype of that cluster.

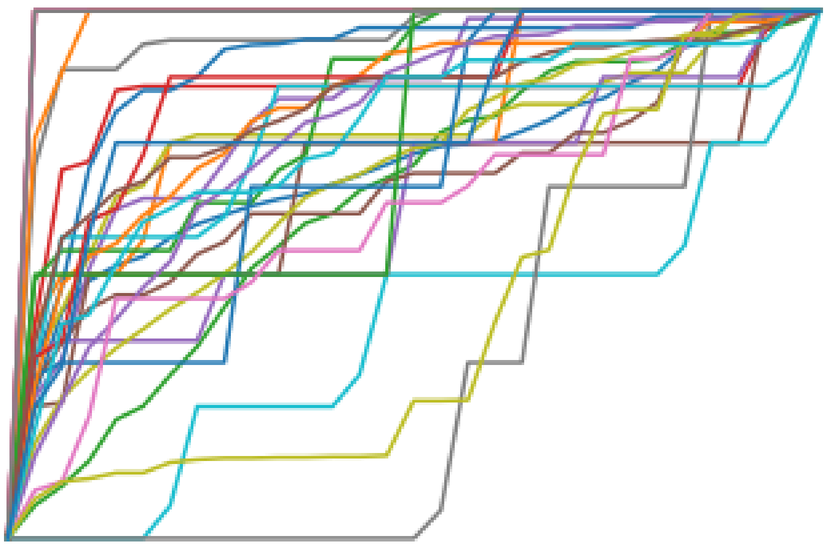

Fig 4.2: Centroids for each cluster where K=50

The clustering results are shown in Figure 4.3, Figure 4.4, Figure 4.5 and Figure 4.6. All of these results differs the value of K. Figure 4.3 represents the results of clustering when K=50, Figure 4.4 represents the results of clustering when K=45, Figure 4.5 represents the results of clustering when K=40 and Figure 4.6 represents the results of clustering when K=10. Each image in training set is assigned a shape prototype which converts the unsupervised data into supervised data. Then the supervised training dataset is used to train a classifier that only consider the social features of an image to predict the sequence shape s_{shape} by assigning the corresponding prototype. Different classifiers were used to evaluate the best one with different combination of features. The best results in our approach are obtained by using RNDF classifier with different combination of social features. It can be summed up that in our proposed model, a given image and its social features are enough to predict the sequence shape by using RNDF classifier.

Fig. 4.3: The Prototypes Obtained after clustering popularity sequences into 50 clusters.

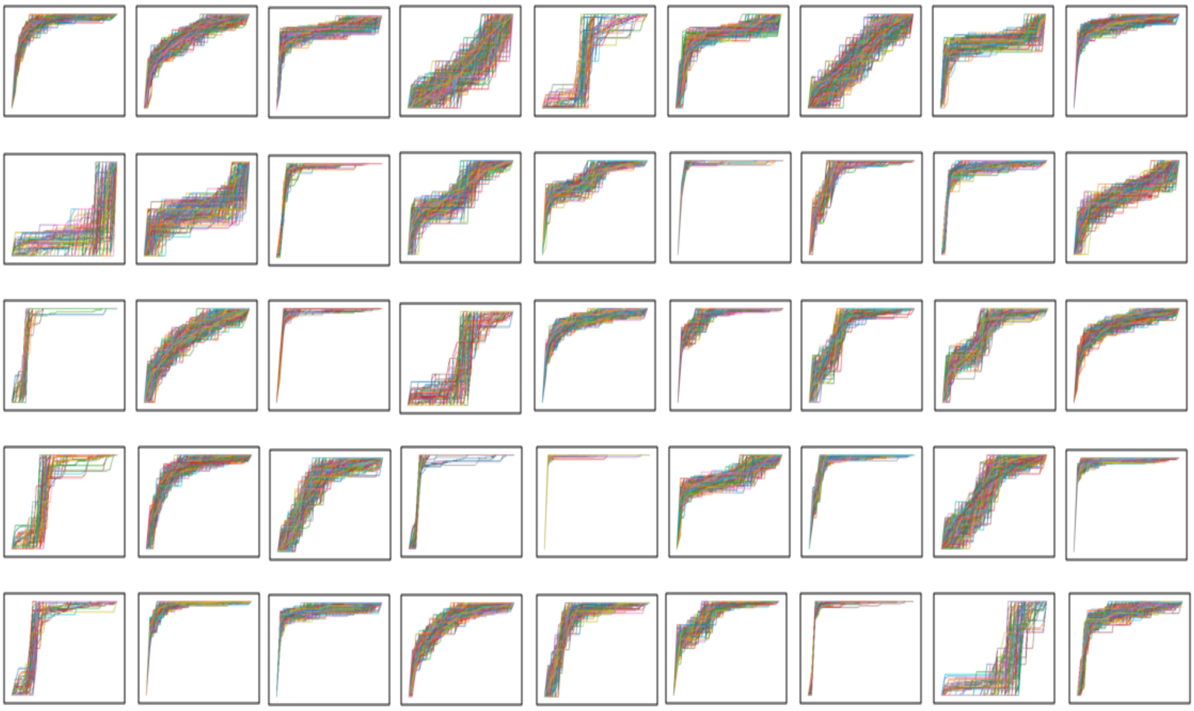

Fig. 4.4: The Prototypes Obtained after clustering popularity sequences into 45 clusters.

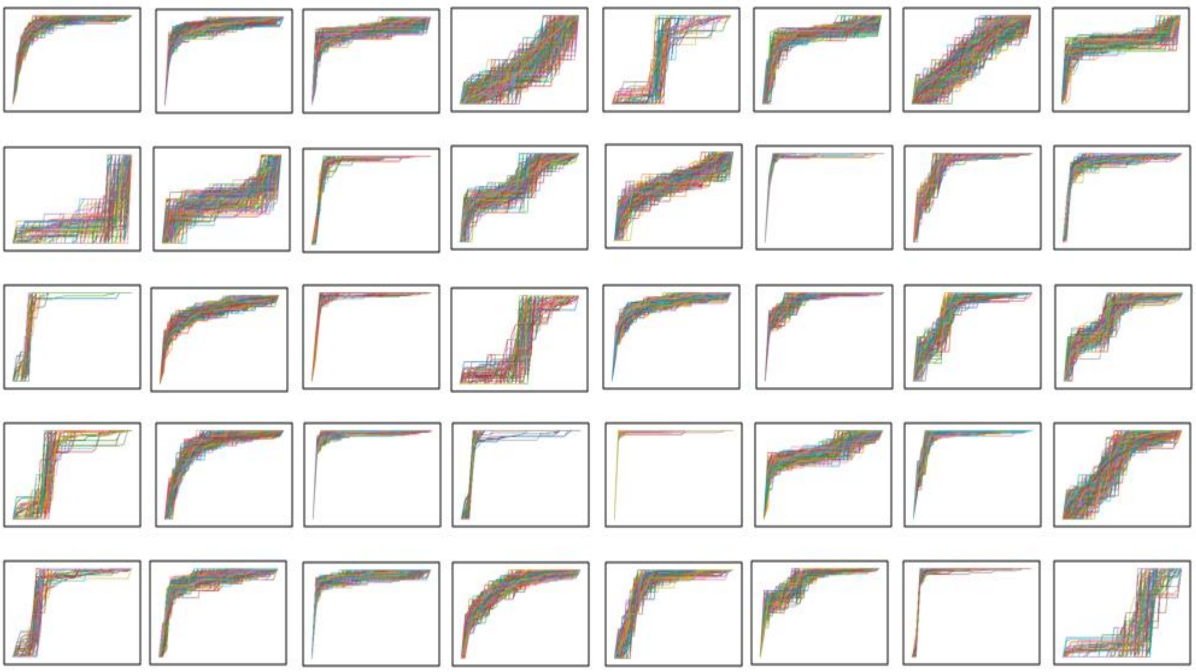

Fig. 4.5: The Prototypes Obtained after clustering popularity sequences into 40 clusters.

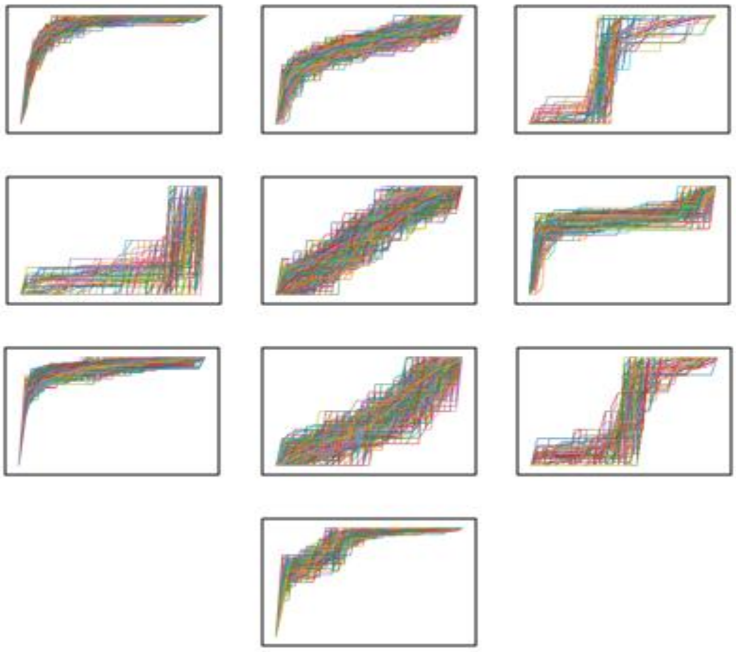

Fig. 4.6: The Prototypes Obtained after clustering popularity sequences into 10 clusters.

5.2. Sequence Scale Estimation

For the estimation of scale value, different regression algorithms are evaluated to find the best one among them. Following Regression Algorithms have been evaluated: SVR RBF, SVR Linear, SVR Sigmoid, SVR Poly, Linear Regression, RF Regression, Local Polynomial Regression, Lasso Regression, Logistic Regression, ElasticNet Regression and Ridge Regression. Results obtained from these regressors in terms of Root Mean Squared Error (RMSE) and spearman's correlation between the predicted values and the ground truth values. SVR RBF provides the best regression results with maximum spearman's correlation and minimum RMSE value. Different sets of features are considered in the experiment. The model is trained to predict the popularity by giving every single social feature and different group of features. The model is also trained by using the recommended features given in [21] and [32]. Afterward, the features are reduced using linear transformation techniques LDA (Supervised) and PCA (un-supervised). These algorithms reduces the features dimension to the given no of features, LDA gives very promising results for our model by reducing dimension from 19 to 6. We have also used different wrapper methods to find out the best subset of the features to predict the popularity. Following methods are used: Forward Selection Method, Backward Selection Method and Bi-directional Method.

Table 4.1 shows the results obtained by the model on each step of methodology. It also shows the best result obtained after running the complete model using best features.

Table 4.1: Results for the prediction at each step of prediction

	tRMSE	Spearman Correlation
Scale	0.225	-
Shape	-	0.76
Scale + Shape	8.03	0.76

4.3. Algorithms Used

There are different algorithms used to design this model. In this part those algorithms are briefly explained.

4.3.1 Support Vector Regression (SVR)

In a machine learning and data science Support Vector machines SVM are very common but SVR is a bit different from SVM. SVR is a regression algorithm. It can be used for prediction where dealing is with continuous value, not with fixed number of classes like in SVM. SVR is used to fix the error rate by using a threshold value rather than just minimizing the error rate like in simple regression.

4.3.2 Linear Discriminant analysis (LDA)

One of the dimensionality reduction technique is LDA. In machine learning it is used as a pre-processing step, and afterward the dimensionally reduced data is used for processing. Purpose of LDA is to reduce the higher dimensional data into lower dimensionality to reduce the resource usage and dimensionality cost. It is a competitive machine learning technique that can perform supervised dimensionality reduction task by considering the labels.

4.3.3 Principal Component Analysis (PCA)

Another technique used for dimensionality reduction is PCA. It is a very simple method that uses linear algebra matrix operation to calculate the original data projection into fewer number of dimensions. It can project m-features into the fewer or desired number of dimensions. It is an un-supervised dimensionality reduction technique, it means it do not consider labels while reducing the dimensions.

4.3.4 Wrapper Methods

Wrapper methods are used where dataset consist of so many features and we do not know which features are to be used. In order to perform machine learning, we need to do feature selection from a larger range of features and this is a very important task. There are some less-significant and irrelevant features in the dataset due to which we have to face some issues like model complexity, time complexity, and dumb model. It is needed to identify the irrelevant features so that the model can work on the best features set to increase the accuracy. Feature selection is a must and crucial task in machine learning. The feature selection process in wrapper methods are based on a greedy search approach, it evaluates all the possible combination of features to find the best one. Most commonly used wrapper methods are:

- Forward Selection
- Backward Elimination
- Bi-directional Elimination (Stepwise Selection)

All of these methods work to find best feature set but use different techniques to achieve this task. The best features subset provided by Forward Selection, Backward Elimination and Bi-directional Elimination include features:

- DateCrwal
- DateTaken
- DatePosted
- AvgGroupsMemb

The best features subset selected by Sequential Feature Selector include features:

- DatePosted
- DateTaken
- DateCrawl
- Size
- Title
- NumGroups
- AvgGroupsMemb
- Contacts
- PhotoCounts
- GroupAvgMembers
- GroupAvgPictures

CHAPTER 5: EXPERIMENTAL RESULTS

5.1. Performance Measures

To measure the performance of the model we have used two performance metrics: Root Mean Squared Error (RMSE) and Spearman's correlation value. Higher Spearman's value indicates that there is higher correlation among the features that are used for prediction. When the Average RMSE value is evaluated for the model, a few larger values can skew the overall average RMSE value. Popularity prediction is very sensitive to the errors related to the sequence scale value. If the ground truth scale values for each cluster are observed, it can be seen that there are values in some clusters that are very larger than the most of the values in the same cluster. This issue of skewed RMSE value is solved by considering the trimmed RMSE value as the performance parameter. It provides the robustness to our model by eliminating the outlier values. We have considered tRMSE 0.25 which means 25% truncated RMSE value. The truncated RMSE is also called inter-quartile mean which means that the same percentage is discarded from the lower and higher tails of the distribution. This means that 25% of both worst and best values are truncated while calculating the average RMSE value. Equation 8 and 9 shows the calculation of spearman correlation and tRMSE values.

$$p = 1 - \frac{6\sum {d_i}^2}{n(n^2-1)} \qquad (8)$$

$$tRMSE = \frac{2}{n}\sum_{\frac{n}{4}+1}^{\frac{3n}{4}} x_i \qquad (9)$$

5.2. Results and Comparisons

Using the given social features for images, this model predicts the sequence shape by using RNDF classifier and assign the corresponding shape prototype s'_{shape}. Then an SVR is used for the estimation of sequence scale sˆscale. The final sequence s' is obtained by multiplying s'_{shape} and s'_{scale} as given in equation (4). This process is repeated multiple times with different number of clusters and different combination of features. In this section the evaluation is performed on the proposed model for the task of IPP for 30 days.

Here the empirical analysis and comparison is done on the results. As discussed results are measured by two parameters tRMSE and Spearman correlation. The Spearman correlation is measured for the SVR that perform scale estimation by using single features and combinational features. tRMSE value is measured for the complete model after prediction with the ground truth values. We have calculated these two parameters for every single feature, the results point out that the user features gives the best results in IPP. It means that the popularity of an image depends on the capability of the user and the potential of the audience of the platform where image is uploaded.

For the estimation of scale value, different regression algorithms are evaluated to find the best one among them. Following Regression Algorithms have been evaluated: SVR RBF, SVR Linear, SVR Sigmoid, SVR Poly, Linear Regression, RF Regression, Local Polynomial Regression, Lasso Regression, Logistic Regression, ElasticNet Regression and Ridge Regression. Results obtained from these regressors in terms of Root Mean Squared Error (RMSE) and spearman's correlation between the predicted values and the ground truth values. Separate results for RMSE values with respect to each Regression algorithm, are shown in Figure 5.1. Results of each regression in the form of Spearman correlation is shown in Figure 5.2. The combined results of RMSE and Spearman correlation for each regression algorithms are shown in Figure 5.3. As we need to select the best results which means maximum value of spearman correlation and minimum value of tRMSE value. SVR RBF provides the best regression results with maximum spearman's correlation and minimum RMSE value.

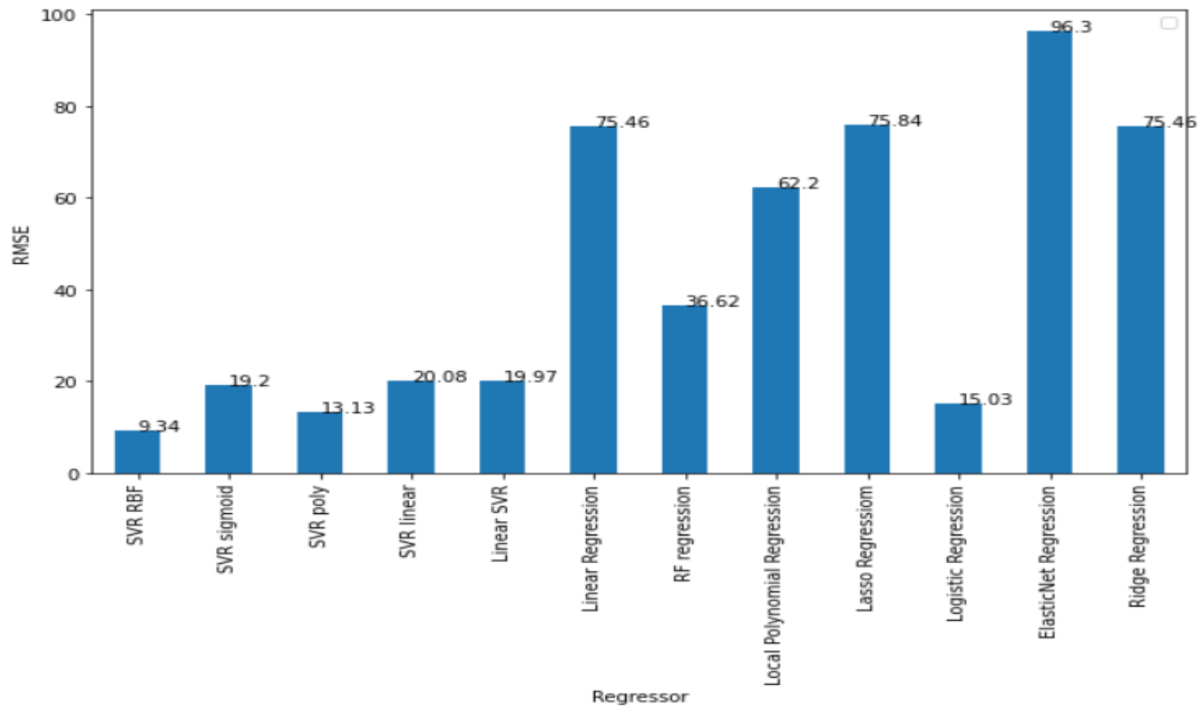

Fig 5.1: RMSE values for each regression algorithm

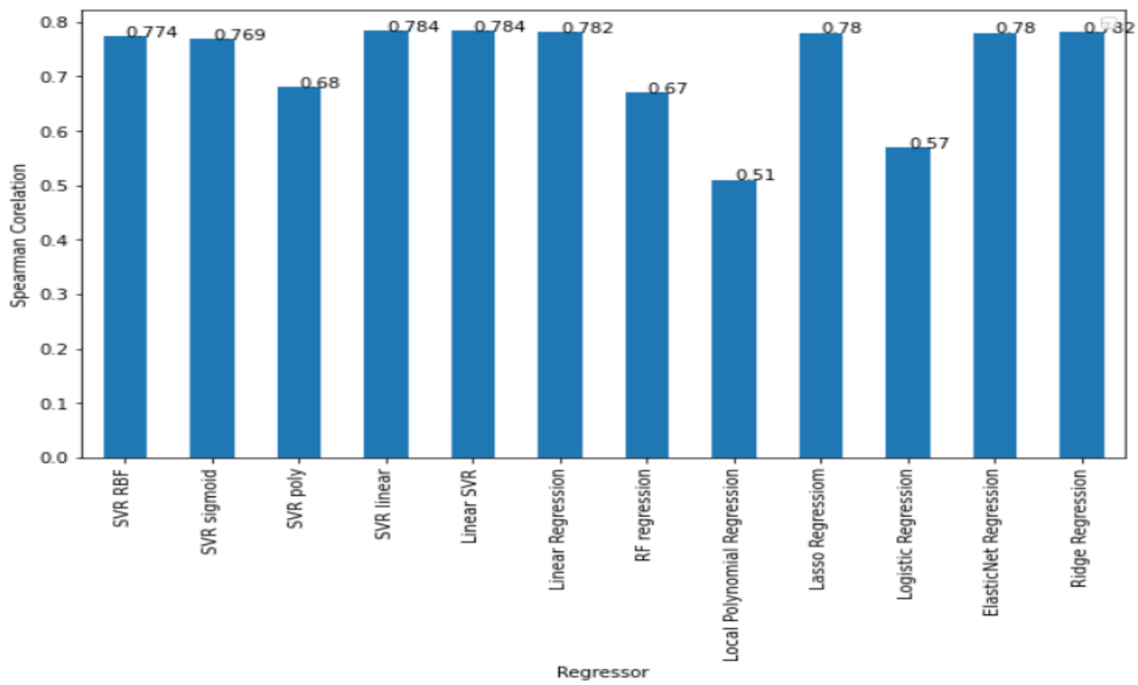

Fig 5.2: Spearman Correlation values for each regression algorithm

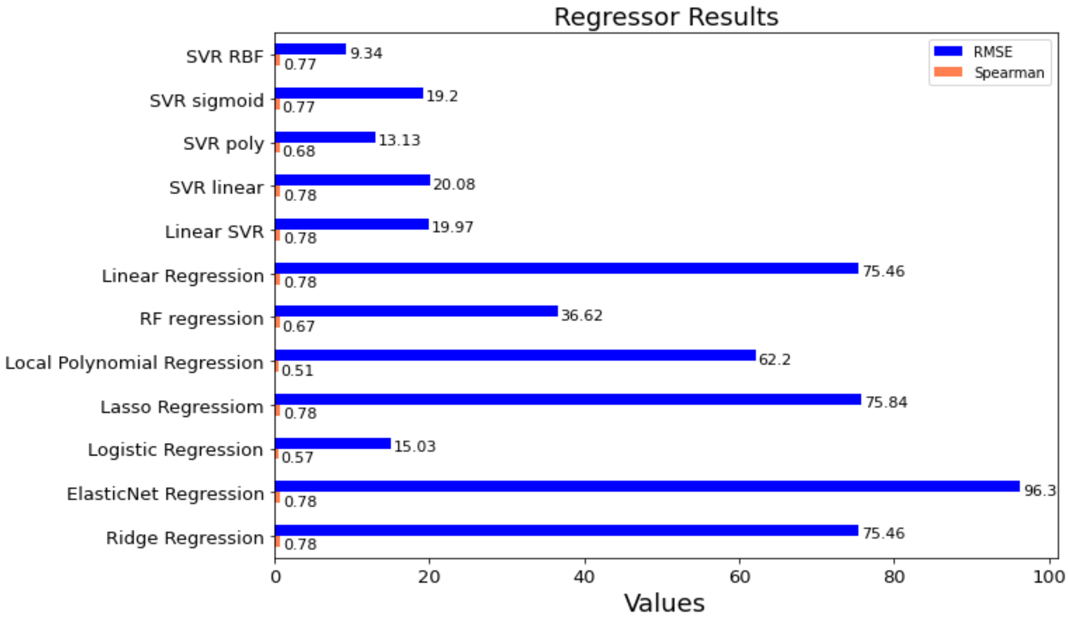

Fig. 5.3: This bar chart show the results of different regressors. The results are given in two parameters RMSE and Spearman correlation.

The results in Table 5.1 depicts that the user feature MeanViews gives the highest Spearman correlation value and minimum tRMSE value. It can be clearly seen in the table that features NumGroups, AvgGroupMemb, AvgGroupPhotos, Contacts and GroupCount gives good results. It means that the context features related to an image such as user features and photo features etc outperforms all other features related to image content. That is why we have used the combination of these features to predict the popularity score and these features are recommended as best in [32]. Row 23 of the Table 6.1 shows the results for the features that are combined according to the recommendation in [21]. This combination of features do not give results for our model.

We have also evaluated our model by find the best subset of features, using wrapper methods: Forward Wrapper, Backward Wrapper and Bi-directional Wrapper. Results obtained from the recommended feature subsets given by these wrapper methods do not differ much from each other in terms of tRMSE and Spearman.

We have used linear transformation techniques on our set of features to reduce the dimension of the features. For this two techniques Principal component analysis (PCA) and Linear Discriminant Analysis (LDA) are used. As PCA is an unsupervised algorithm the results obtained after reducing dimensions using this technique are not satisfying and almost same for all dimensionalities. On the other hand LDA is a supervised algorithm and it gives very promising results for our proposed model. LDA reduced the dimensionality with respect to the labels, in our case when we reduce the dimensionality of features from 19 to 6 our model gives best results; higher Spearman correlation value of 0.76 and lowest tRMSE value of 8.03.

The final prediction results are shown in Figure 5.4. The plots in the figure shows the popularity prediction score for the span of 30 days for four different Flickr images. Blue lines show the popularity score predicted by our model and the orange lines are the ground truth popularity scores of the images. It can be seen that our proposed model gives very good prediction results.

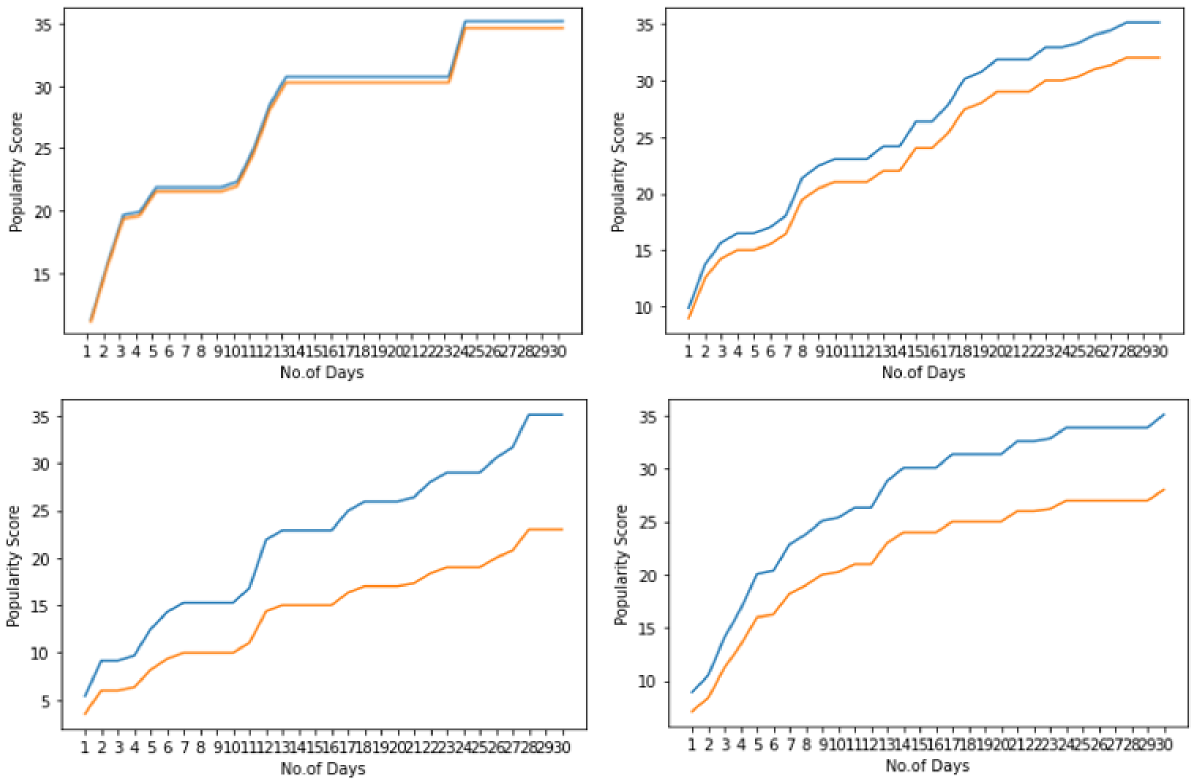

Fig. 5.4: The prediction results for the model: In the plots the predicted sequence scores are plotted with respect to the ground truth sequence scores. Our model gives very good prediction results.

Table 5.1: Results of popularity prediction for the span of 30 days

S.No	Feature	Spearman	RMSE	Trimmed RMSE
1	DatePosted	0.064	135.885	12.75
2	DateTaken	0.064	135.89	12.75
3	DateCrawl	0.57	135.889	12.75
4	Size	0.57	135.7	13.19
5	Title	0.13	135.71	14.08
6	Description	0.047	135.84	13.64
7	NumSets	0.25	133.51	13.42
8	NumGroups	0.41	122.82	10.19
9	AvgGroupsMemb	0.41	125.06	10.9
10	AvgGroupPhotos	0.41	125.4	10.95
11	Tags	0.19	134.62	15.04
12	Ispro	0.33	134.11	13.4
13	HasStats	0.36	133.86	13.22
14	Contacts	0.599	124.68	10.44
15	PhotoCount	0.182	135.85	13.43
16	MeanViews	0.77	106.9	9.52
17	GroupsCount	0.61	123.99	11.29
18	GroupsAvgMembers	0.56	131.1	15.32
19	GroupsAvgPictures	0.56	130.9	16.08
20	concat(8,9,10,14,16,17)	0.76	125.4	9.06
21	concat(14,16,17)	0.76	109.99	9.83
22	concat(8,9,10)	0.41	124.42	10.23
23	concat(5,6,8,9,10,11,12,14,15,16,17)	0.14	126.2	12.9
24	PCA 18-2	0.114	135.85	13.89
25	LDA 18	0.705	111.96	12.16
26	LDA 16	0.704	111.61	12.25
27	LDA 14	0.703	110.5	12.13
28	LDA 12	0.71	109.42	9.33
29	LDA 10	0.69	108.24	9.28
30	LDA 8	0.72	106.31	8.71
31	**LDA 6**	**0.76**	**104.48**	**8.03**
32	LDA 4	0.75	103.6	8.06
33	LDA 2	0.72	104.26	9.08
34	Forward Wrapper	0.061	135.88	10.43
35	Backward Wrapper	0.061	135.88	10.59
36	Bi-directional Wrapper	0.061	135.88	10.43
37	Built-in (Sequential Feature Selector())	0.061	135.88	10.43

By visualizing the results it can be clearly seen that the larger variation in scale values prediction can affect the overall prediction badly. It can be seen in Figure 5.5. The image/video context features related to user and text outperforms all other visual features of the image/video. We can clearly conclude that some features have strong correlation with the popularity of content on social media.

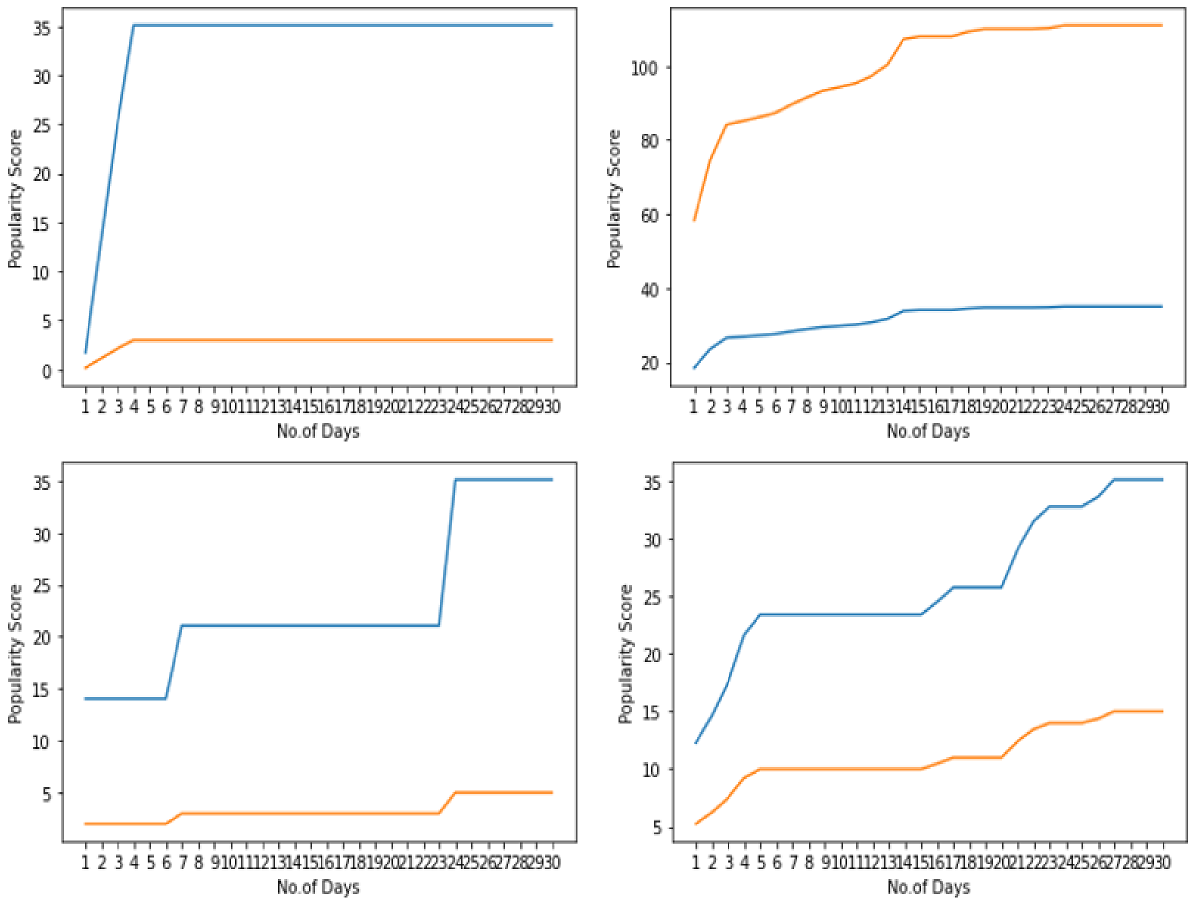

Fig. 5.5: The sensitivity of the model to the sequence scale value is shown, wrong prediction of the scale value can result in very bad prediction.

CHAPTER 6: CONCLUSION & FUTURE WORK

6.1. Conclusion

A lot of the research work in past have been done to address the problem of IPP, but most of them have worked to predict a single popularity score for social content. This work predicts the popularity score for the temporal evolution of the content. In this paper, a model is designed for the challenging task of image popularity prediction for the analysis of social platforms. This work discusses the experimental results of the methodology that is used for the task of prediction of popularity for the span of 30 days. Due to the independence of sequence shape and sequence scale different machine learning techniques can be used to predict the values. An ablation analysis is done by using different features individually and in combination to other features to check their effect on the prediction accuracy. The predicted popularity scores are compared to ground truth popularity scores. This work concludes that the proposed method predicts accurately in case of sequence shape but it is very sensitive to the sequence scale value. A larger variation in scale value prediction can affect the overall prediction badly. The image/video context features related to user and text outperforms all other visual features of the image/video. We can clearly conclude that some features have strong correlation with the popularity of content on social media.

6.2. Contribution

- An in-depth investigation of the factors and features which effects the level of engagement of an image/video among people.
- It addresses the challenging task of IPP that can help in many applications related to advertisement campaigns, recommendation systems, social media marketing and many more.
- A comparative study is performed on the problem with the aim to learn the effect of different type of information on popularity score.
- A system is proposed to design an application to support information diffusion. The system will provide the forecast for the popularity score of image/video for the next 30 days after uploading the content.

6.3. Future Work

There is need for further investigation to address some questions related to image popularity over time and some challenging problems in this domain, which are engaging the attention of ample amount of researchers. Conventional machine learning algorithms are used, different features are manually selected as well as selected by using different techniques, and then a classifier is trained on them. Many other techniques can be evaluated to get better results. Extracting and choosing the optimum feature set according to the problem is a crucial task, there can be some other features that can provide more accurate popularity score for an image. Also many other ensembles can be implemented using combination of different features. New and more efficient model could be designed with a little effort.

REFERENCES

[1] Figueiredo, F., Almeida, J. M., Gonc¸alves, M. A., Benevenuto, F. (2014). On the dynamics of social media popularity: A YouTube case study. ACM Transactions on Internet Technology (TOIT), 14(4), 1-23.

[2] McParlane, P. J., Moshfeghi, Y., Jose, J. M. (2014, April). " Nobody comes here anymore, it's too crowded"; Predicting Image Popularity on Flickr. In Proceedings of International Conference on Multimedia Retrieval (pp. 385-391).

[3] Totti, L. C., Costa, F. A., Avila, S., Valle, E., Meira Jr, W., Almeida, V. (2014, June). The impact of visual attributes on online image diffusion. In Proceedings of the 2014 ACM conference on Web science (pp. 42-51).

[4] Khosla, A. (2014). K., Das Sarma, A., Hamid, R.: What makes an image popular. In Proceedings of the 23rd international conference on World wide web (pp. 867-876).

[5] Gelli, F., Uricchio, T., Bertini, M., Del Bimbo, A., Chang, S. F. (2015, October). Image popularity prediction in social media using sentiment and context features. In Proceedings of the 23rd ACM international conference on Multimedia (pp. 907-910).

[6] Aloufi, S., Zhu, S., El Saddik, A. (2017). On the prediction of flickr image popularity by analyzing heterogeneous social sensory data. Sensors, 17(3), 631.

[7] Hu, J., Yamasaki, T., Aizawa, K. (2016, May). Multimodal learning for image popularity prediction on social media. In 2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCETW) (pp. 1-2). IEEE.

[8] Trzcinski, T., Rokita, P. (2017). Predicting popularity of online videos ´ using support vector regression. IEEE Transactions on Multimedia, 19(11), 2561-2570.

[9] Keneshloo, Y., Wang, S., Han, E. H., Ramakrishnan, N. (2016, June). Predicting the popularity of news articles. In Proceedings of the 2016 SIAM International Conference on Data Mining (pp. 441-449). Society for Industrial and Applied Mathematics.

[10] Mazloom, M., Hendriks, B., Worring, M. (2017, October). Multimodal context-aware recommender for post popularity prediction in social media. In Proceedings of the on Thematic Workshops of ACM Multimedia 2017 (pp. 236-244).

[11] Wu, B., Cheng, W. H., Zhang, Y., Huang, Q., Li, J., Mei, T. (2017). Sequential prediction of social media popularity with deep temporal context networks. arXiv preprint arXiv:1712.04443.

[12] Fernandes, K., Vinagre, P., Cortez, P. (2015, September). A proactive intelligent decision support system for predicting the popularity of online news. In Portuguese Conference on Artificial Intelligence (pp. 535-546). Springer, Cham.

[13] Nwana, A. O., Avestimehr, S., Chen, T. (2013, December). A latent social approach to youtube popularity prediction. In 2013 IEEE Global Communications Conference (GLOBECOM) (pp. 3138- 3144). IEEE.

[14] Mazloom, M., Pappi, I., Worring, M. (2018, February). Category specific post popularity prediction. In International Conference on Multimedia Modeling (pp. 594-607). Springer, Cham.

[15] Jheng, S. H., Li, C. T., Chen, H. L., Shan, M. K. (2013, September). Popularity Prediction of Social Multimedia Based on Concept Drift. In 2013 International Conference on Social Computing (pp. 821-826). IEEE.

[16] De, S., Maity, A., Goel, V., Shitole, S., Bhattacharya, A. (2017, April). Predicting the popularity of instagram posts for a lifestyle magazine using deep learning. In 2017 2nd International Conference on Communication Systems, Computing and IT Applications (CSCITA) (pp. 174-177). IEEE.

[17] Mao, Y., Monahan, A. (2018). Linear and nonlinear regression prediction of surface wind components. Climate Dynamics, 51(9-10), 3291-3309.

[18] Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. American scientist, 89(4), 344-350.

[19] Borth, D., Ji, R., Chen, T., Breuel, T., Chang, S. F. (2013, October). Large-scale visual sentiment ontology and detectors using adjective noun pairs. In Proceedings of the 21st ACM international conference on Multimedia (pp. 223-232).

[20] Chen, T., Borth, D., Darrell, T., Chang, S. F. (2014). Deepsentibank: Visual sentiment concept classification with deep convolutional neural networks. arXiv preprint arXiv:1410.8586.

[21] Chen, T., Yu, F. X., Chen, J., Cui, Y., Chen, Y. Y., Chang, S. F. (2014, November). Object-based visual sentiment concept analysis and application. In Proceedings of the 22nd ACM international conference on Multimedia (pp. 367-376).

[22] Yu, H., Bai, X. F., Huang, C., Qi, H. (2015). Prediction of users retweet times in social network. International Journal of Multimedia and Ubiquitous Engineering, 10(5), 315-322.

[23] Hong, L., Dan, O., Davison, B. D. (2011, March). Predicting popular messages in twitter. In Proceedings of the 20th international conference companion on World Wide Web (pp. 57-58).
[24] Zaman, T. R., Herbrich, R., Van Gael, J., Stern, D. (2010, December). Predicting information spreading in twitter. In Workshop on computational social science and the wisdom of crowds, nips (Vol. 104, No. 45, pp. 17599-601). Citeseer.
[25] Can, E. F., Oktay, H., Manmatha, R. (2013, October). Predicting retweet count using visual cues. In Proceedings of the 22nd ACM international conference on Information Knowledge Management (pp. 1481-1484).
[26] Fiolet, E. (2014). Analyzing image popularity on a social media platform. Unpublished Master Thesis), University of Amsterdam. Retrieved January, 4, 2015.
[27] Cappallo, S., Mensink, T., Snoek, C. G. (2015, June). Latent factors of visual popularity prediction. In Proceedings of the 5th ACM on International Conference on Multimedia Retrieval (pp. 195-202).
[28] Yamaguchi, K., Berg, T. L., Ortiz, L. E. (2014, November). Chic or social: Visual popularity analysis in online fashion networks. In Proceedings of the 22nd ACM international conference on Multimedia (pp. 773-776).
[29] Niu, X., Li, L., Mei, T., Shen, J., Xu, K. (2012, July). Predicting image popularity in an incomplete social media community by a weighted bi-partite graph. In 2012 IEEE International Conference on Multimedia and Expo (pp. 735-740). IEEE.
[30] Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A. (2014). Return of the devil in the details: Delving deep into convolutional nets. arXiv preprint arXiv:1405.3531.
[31] Ortis, A., Farinella, G. M., Battiato, S. (2019). Predicting social image popularity dynamics at time zero. IEEE Access, 7, 171691- 171706.
[32] Zohourian, A., Sajedi, H., Yavary, A. (2018, April). Popularity prediction of images and videos on Instagram. In 2018 4th International Conference on Web Research (ICWR) (pp. 111-117). IEEE.
[33] Almgren, K., Lee, J., Kim, M. (2016, October). Prediction of image popularity over time on social media networks. In 2016 Annual Connecticut Conference on Industrial Electronics, Technology Automation (CT-IETA) (pp. 1-6). IEEE.
[34] Meghawat, M., Yadav, S., Mahata, D., Yin, Y., Shah, R. R., Zimmermann, R. (2018, April). A multimodal approach to predict social media popularity. In 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) (pp. 190-195). IEEE.

[35] iplab.dmi.unict.it. Retrieved November 26, 2020 from https://iplab.dmi.unict.it/popularitydynamics

[36] yfcc100m.org. Retrieved November 26, 2020 from http://projects.dfki.uni-kl.de/yfcc100m/

[37] kaggle.com. Retrieved November 26, 2020 from https://www.kaggle.com/hsankesara/flickr-image-dataset

[38] opensource.google. Retrieved November 26, 2020 from https://storage.googleapis.com/openimages/web/download.htm

[39] Jgmediastrategies.com. Retrieved July 9, 2021 from https://www.jgmediastrategies.com/our-impact.

[40] online.maryville.edu. Retrieved July 9, 2021 from https://online.maryville.edu/blog/evolution-social-media/