

CONVERGENCE

International Journal of ICTAK

(A Multidisciplinary Journal of Engineering, Technology & Employability)



ICT ACADEMY OF KERALA

(A Govt. of India supported, Govt. of Kerala partnered Social Enterprise)



EDITORIAL BOARD

Chief Patron

Dr. Sandhta Chintala

Vice President - NASSCOM &
Executive Director - Sector Skills Council, NASSCOM

Chief Editor

Dr. Manoj AS

Head - Corporate Operating Unit,
ICT Academy of Kerala

Editorial Advisory Board

Dr. Rajasree M.S

Vice Chancellor
APJ Abdul Kalam Technological University
Thiruvananthapuram.

Prof. Mithileshwar Jha

Professor - Marketing (Retired)
Indian Institute of Management, Bangalore

Dr. A. Damodaran

Chairperson, economics & Social Science Area,
Chair Professor IPR Chair on IP Management (MHRD),
Indian Institute of Management, Bangalore

Prof. L.S Murthy

Professor - Production & Operations Management,
Indian Institute of Management, Bangalore

Dr. Prahlad Vadakkepat

Associate Professor,
Electrical & Computer Engineering
National University of Singapore (NUS)

INTRODUCTION

Pedro Raymond once said “Not everything you think of is worth being published. But everything you think of is worth being written.” ICT Academy of Kerala is delighted to present before our erudite writers, scholars and all stakeholders to our 5th edition of the International Journal - “Convergence” - A Multidisciplinary Journal which aims to provide novel perspectives and ideas on Engineering, Technology & Employability where academicians, research scholars, industry experts and professionals can contribute their innovative and creative ideas in the form of research articles, case studies and book reviews.

The articles included in this journals were predominantly from the authors who have presented their papers on the ICTAK International Conclave - ICSET 2020 on the theme “IR 4.0 in action”. Industry 4.0 is a vision that evolved from an initiative to make the German manufacturing industry more competitive ('Industrie4.0') to a globally adopted term. Industry 4.0 is often used interchangeably with the notion of the fourth industrial revolution. It is characterized by, among others, 1) even more automation than in the third industrial revolution, 2) the bridging of the physical and digital world through cyber-physical systems, enabled by Industrial IoT, 3) a shift from a central industrial control system to one where smart products define the production steps, 4) closed-loop data models and control systems and 4) personalization/customization of products.

On the above contexts, there were plethora of discussions and deliberations with regard to the theme and there were paper presentations by faculty members and research scholars on the above theme. The selected papers were included in this edition of Convergence, which will give insights, informations and cognizance with regard to the theme.

We wish you a fruitful and a very informative reading!

Dr. Manoj AS

Chief Editor - Convergence

COVID-19 - DATA ANALYSIS

Dr. Tony Thomas, TonyT123@gmail.com

If for livelihood, we have to traverse through an unknown terrain, where no one has been before, like a swamp that is unfamiliar to us, uncharted, covered with bushes with dangerous landmines? As a first step we will turn to experts and ask what we should we do. Some of the experts may say, "Stay home". Others will say, "What's the big deal—there's nothing to fear".

We can stay home as some of the Experts say and starve to death or run out recklessly and die in the swamps based on the others. If we don't go, we will starve to death along with the rest of the family. So, we assume the truth is somewhere in the middle, take the chance and venture out with enough preparation, each step will be taken gently with great care. If we suspect danger with any step, at once we will adjust, correct and move on step by step. This is how we face any crisis. Before we take each step, we look at the data, and make appropriate decisions. Data helps us make unbiased views and rational decisions. Constant feedback loop ensures we keep adjusting or pivoting as required.

Coronavirus related COVID-19 began in China in late 2019 and has spread to many countries by January. The disease can spread from an average of 2.2 people per patient, with a spread interval of 6 days. Some of the worst impacted countries have passed through the peak already. We can use the knowledge gathered from many countries as data points to make decisions.

As we have few countries that already passed through the peak and are beyond the danger phase, we have substantial amount of data that can be used for decision making. We should analyse not just the impact of the risk, but also the probability of the risk as well. Cognitive biases can lead to poor decisions, so they have to be explicitly be discounted. In the context of Coronavirus, 9-11 terror attack in New York, and similar crises, fear can cause many biases. We also have to be mindful of order of magnitudes and cause and effects.

If we go by the number predictions, a seed, an index patient on January 1 would have caused over 6 million cases by mid-march. That is clearly not the case. Most draconian interventions are based on such numbers. We are probably in an environment where the drastic interventions may have adverse impact on more lives and livelihood than from the pandemic itself.

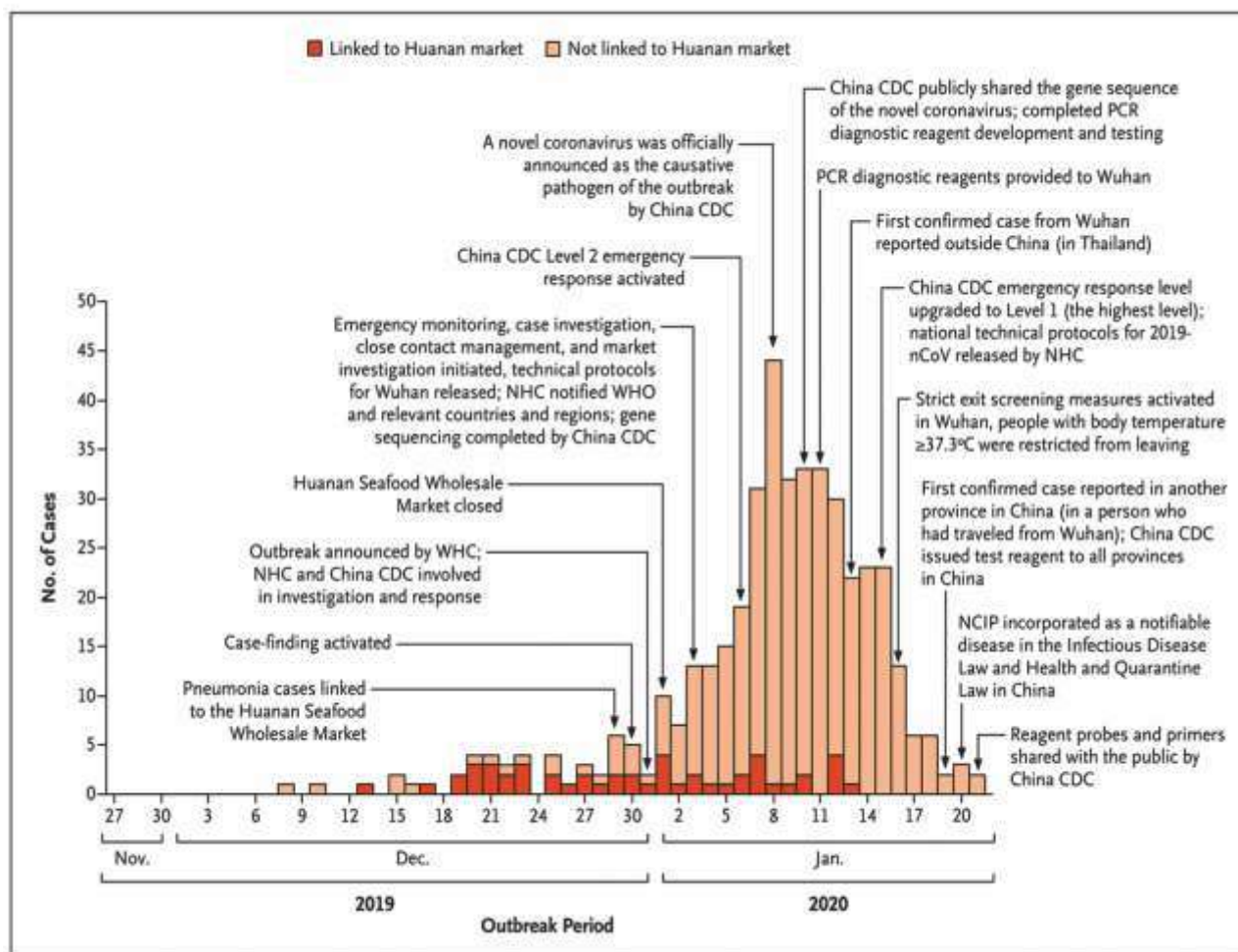
We can look at overall population, disease rate within population, death rate within population, the rate of progression, intervening actions that can cause changes to these etc. We can also look at the emerging trends including 7-day trend, 3-day tend etc. We can also conduct comparative analysis from places with draconian controls, partial controls and no-controls. These trends, normalized and transposed on current situation can give us a reasonable insight on what to expect.

As this is a virus, World Health Organization WHO and similar organizations are publishing a lot of relevant guidelines and data that can help decisions. As per WHO, the only way to prevent the disease is through immunity. Immunity comes from vaccinations, inherently in the body or from exposure to the virus. We can prevent uncontrolled spread through physical distancing and certain hygiene controls. If we identify a case through symptoms or testing, we can follow protocols to quarantine and treat them, use contract tracing to identify all contacts who have been in contact, test them, quarantine them or treat them as required. If they fail to keep the disease in check and in line with the Healthcare capacity, the system can enforce more draconian controls like mandatory lockdowns, etc. However, Lockdowns and other measures can only delay, but not prevent the disease. The treatments available are not for the virus, but for the symptoms, hence treatment will not cure the disease. The old saying, "If a doctor treats your cold virus, it will go away in fourteen days. If you leave it alone, it will go away in two weeks" becomes relevant in this context.

2. Coronavirus and COVID-19

COVID-19 is one of seven types of coronavirus, including the ones that cause severe diseases like Middle East Respiratory Syndrome (MERS) and Sudden Acute Respiratory Syndrome (SARS). The other coronaviruses cause most of the colds that affect us during the year but aren't a serious threat for otherwise healthy people.

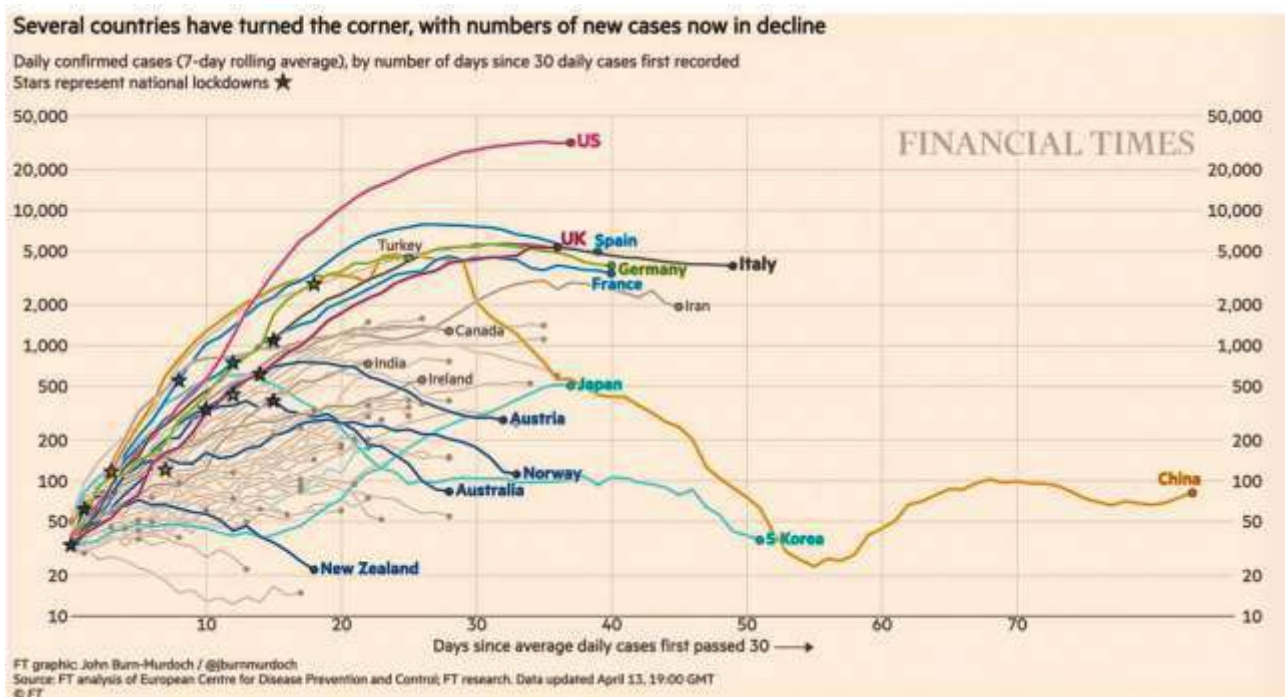
If we look at the below image, we can analyse how the number of cases increased in Wuhan (first epicentre of the outbreak) throughout the outbreak period starting from end of November.



After Wuhan, the virus soon started spreading across the globe. As of April 12, there are 17,83,941 confirmed cases and 109,312 deaths globally.

Kerala reported the first case on January 30, 2020 of three Wuhan returnees. As of April 12, 2020, Kerala

had a total of 376 cases reported, of which 231 patients came from abroad, 18 were domestic travellers and 127 contracted the disease through contact. There were also 2 reported deaths (ages of 68, 69) with pre-existing medical conditions.

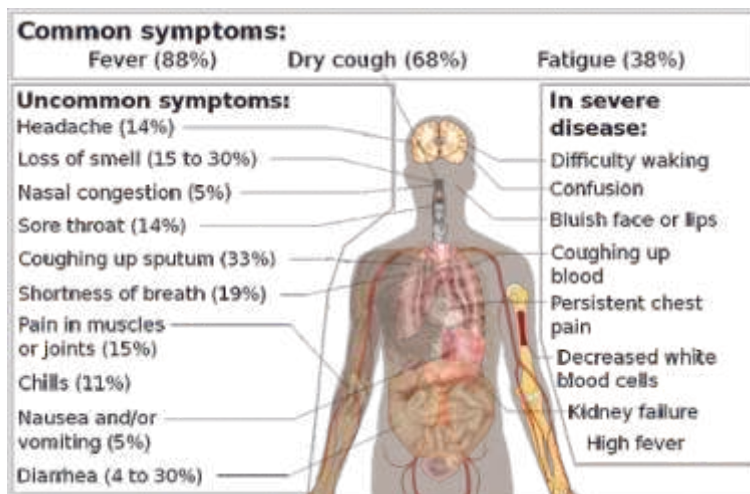


As on April 13, many countries have passed the peak including the worst affected Italy, Spain and UK.

3. Symptoms of Covid-19









Coronaviruses are a large family of viruses which may cause illness in animals or humans. In humans, several coronaviruses are known to cause respiratory infections ranging from the common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). The most recently discovered coronavirus causes the disease which has been named as COVID-19.

Coronavirus disease 2019 (COVID-19), is caused by Severe Acute respiratory Syndrome Coronavirus 2 (SARS-CoV-2). It typically spreads via airborne droplets. These droplets from an infected person's coughs or sneezes contain live coronavirus particles, which can survive for a sufficiently long-time on surfaces. This increases the chances for people to get COVID-19 by touching a contaminated surface and then their face.



The time between exposure and symptom onset is typically around 5 days. It is most contagious during the first 3 days after symptom onset. The WHO currently estimates the incubation period for COVID-19 to range from one to 14 days, most commonly around 5 days (for comparison's sake, the flu's incubation period is about 2 days). Mentioned below are the observed symptoms of the virus -

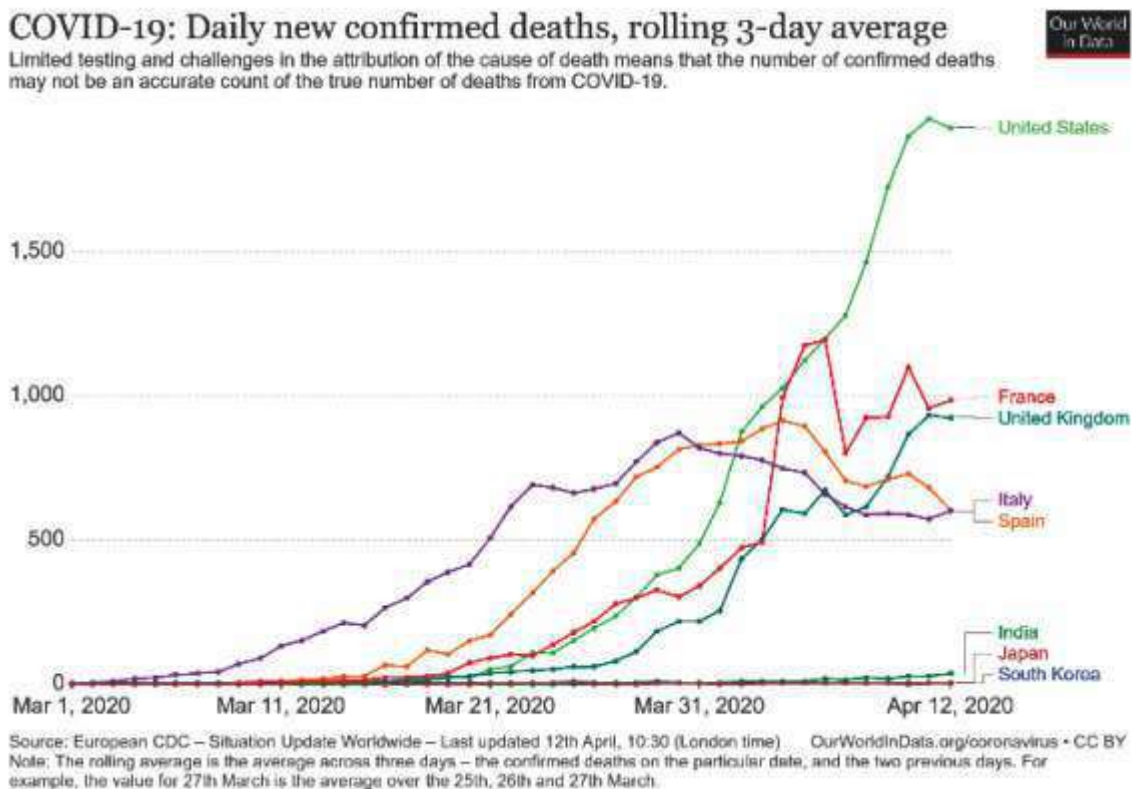
COVID 19 is observed to live on surfaces and anyone who comes in contact with such a contaminated surface and then their face is likely to catch coronavirus (Formite Transmission). The lifespan of the virus on various surfaces is explained in the image below –

How long COVID-19 can live on common surfaces*	
SURFACE	LIFESPAN OF COVID-19
 Air	3 hours 
 Copper	4 hours 
 Cardboard	24 hours 
 Stainless Steel	2–3 days 
 Polypropylene plastic	3 days 

*At 69.8 to 73.4°F (21 to 23 °C) and 40% relative humidity
Source: New England Journal of Medicine
BUSINESS INSIDER

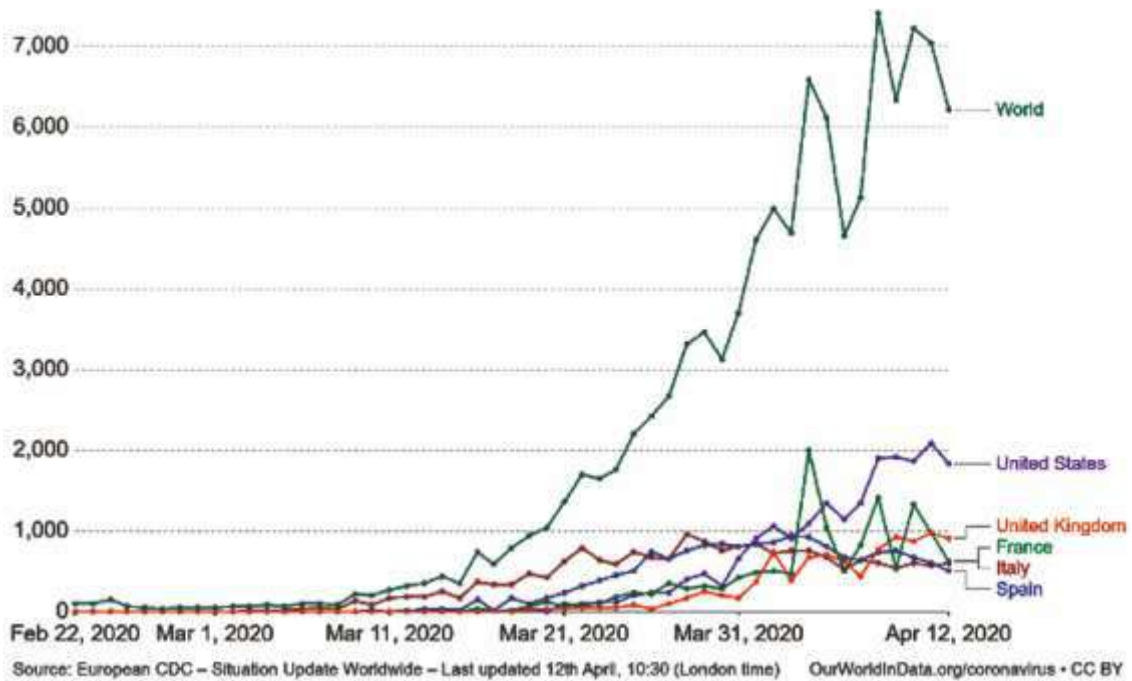
4. Fatality of COVID-19

Let's look at the confirmed COVID-19 cases and the confirmed death rate count around the globe and with other similar viruses.

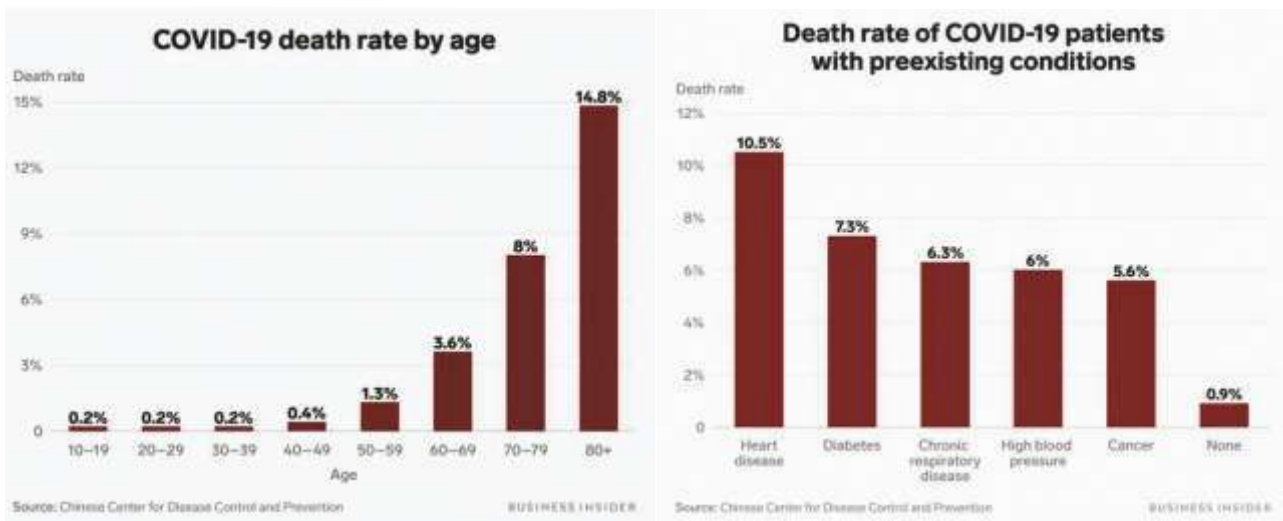


Daily new confirmed COVID-19 deaths

Limited testing and challenges in the attribution of the cause of death means that the number of confirmed deaths may not be an accurate count of the true number of deaths from COVID-19.



If we further analyse the death pattern, the death rate is governed by 2 major factors, Age and Pre-existing condition of the patient. Let's look at the below analysis for the same –



2019 novel coronavirus compared to other major viruses

VIRUS	YEAR IDENTIFIED	CASES	DEATHS	FATALITY RATE	NUMBER OF COUNTRIES
Ebola	1976	33,577	13,562	40.4%	9
Nipah	1998	513	398	77.6%	2
SARS	2002	8,096	774	9.6%	29
MERS*	2012	2,494	858	34.4%	28
COVID-19**	2020	1,677,256	101,732	6%	185

Sources: Johns Hopkins, CDC, World Health Organization, New England Journal of Medicine, Malaysian Journal of Pathology, CGTN

*As of November 2019 **As of April 10, 2020 at 3:25 p.m. EST.

BUSINESS INSIDER

5. Comparison of different viruses

5.1 Comparison between Seasonal Flu and COVID-19

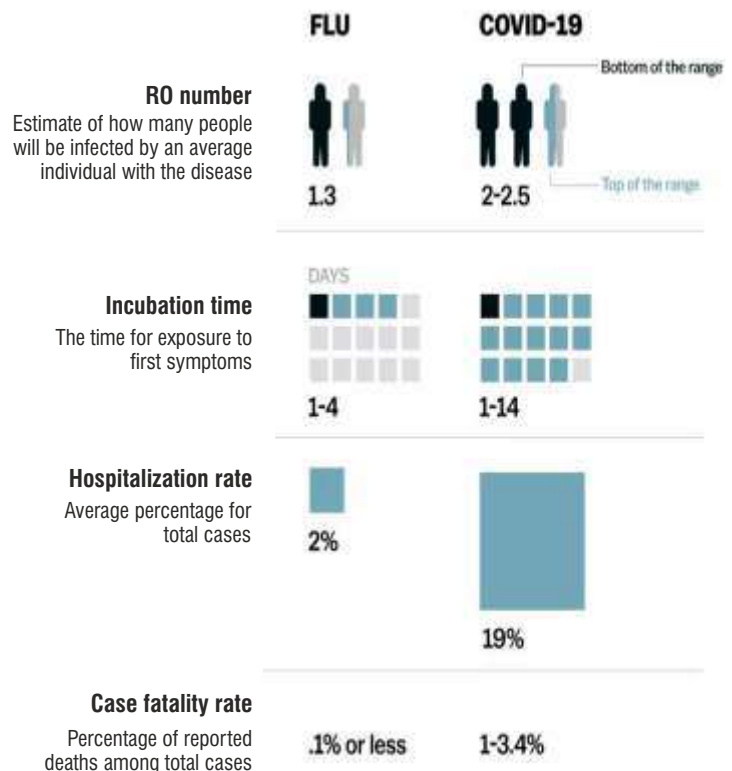
To begin with, let's look at the below image for a quick comparison on Seasonal Flu and COVID-19 -

- As you see in the image above, the R0 number (R naught), the reproduction number is used to describe the intensity of an infectious disease outbreak.
- If R0 is less than 1, the disease will die out in a population, because on average an infectious person will transmit to fewer than one other susceptible person
- On the other hand, if R0 is greater than 1, the disease will spread
- The R0 for Seasonal Flu is 1.3, whereas, the reported R0 for COVID-19 is 2 to 2.5

5.2 Comparison between Influenza and COVID-19

What we see with respect to COVID-19 is that a lot of data is being thrown around, without

How seasonal flu and Covid-19 compare



Sources: CDC, WHO, NCBI

Vox

much substantiation. Many data analysis which creates panic is shared widely and many policy makers use these to make decision purely base on the sensational aspect. Let's take a look at Influenza in the USA which is the closest active disease of volume we have and COVID-19 to dispel some popular urban myths of COVID-19.

Key point to note is that because of inconsistencies, rate of testing, and probably asymptomatic cases, the confirmed cases can be misleading. However, the COVID-19 related deaths can be used with

confidence, as death cannot go unreported.

Based on CDC, COVID-19 and influenza viruses have a similar disease presentation in that they both cause respiratory disease, which represents as a wide range of illness from asymptomatic or mild, through to severe disease and death. and that both viruses are transmitted by contact, droplets and fomites.

As a result, the same public health measures, such as hand hygiene and good respiratory etiquette (coughing into your elbow or into a tissu

Now let's study the patterns of spread for Influenza and COVID-19 in the USA

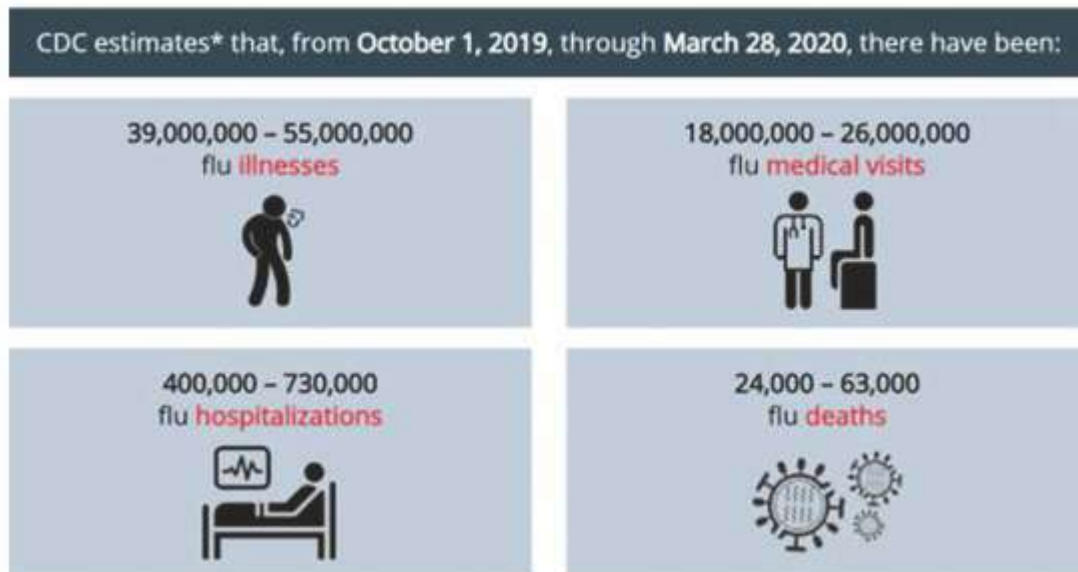
INFLUENZA	COVID-19
The Flu season in USA is tracked from October 1st through May 31st of the next year, ie., 8 months.	COVID-19 was first reported in China in December 2019. The first confirmed case of COVID-19 in USA was reported on January 21st
CDC has published data for 6 months of 2019-2020 flu season	From the date of inception, we have the sick, hospitalization and fatality count for COVID-19 as of April 7
R0 is 1.3	R0 is 2-2.5
The serial interval is 3 days, this means that influenza can spread faster than COVID-19	The serial interval is 5-6 days

From the data we have so far for both Influenza and COVID-19 in USA-

The total population of USA is 330,550,311. As per CDC about 45.3% adults and 62.6% of children get vaccinated for Influenza and the vaccination is effective in about 45%. This means about 73,235,493 Americans are immune to Influenza, or 257,314,818 are vulnerable.

	USA Influenza Vaccination		
	Total Population	Percentage	Count
Adults 18 plus vaccinated	25,53,69,678	45.30%	11,56,82,464
Children vaccinated	7,51,80,633	62.60%	4,70,63,076
Total vaccinated	33,05,50,311	49.23%	16,27,45,540
Flu vaccine effectiveness		45.00%	
Total Protected		22.16%	7,32,35,493
Total Vulnerable		77.84%	25,73,14,818

CDC estimates about 33Mn to 55Mn Americans got sick by influenza, with about 18Mn to 26Mn Medical visits, 400,000 to 730,000 hospital admissions and 24,000 to 63,000 deaths



As of April 8th, USA has 400,728 cases of COVID-19, 41,000 hospitalizations and 12,901 deaths. This gives about 15.16% to 21.37% infection rate for influenza, considering the vulnerable population and, 0.12% infection rate for COVID-19

However, some may argue that the low number of COVID-19 cases is because there are asymptomatic cases and that there is not enough testing to confirm these, so they go unreported. In the case of Diamond

Princess ship (an infected cruise ship which remained isolated due to coronavirus) , it was observed that that it had 18% asymptomatic cases, however, there have been reports that up to 30% can be asymptomatic, which may mean that 30% patients will not request medical help, and may go unreported. So, let's do three 'what if' scenarios to accommodate some of these concerns –

'What if' scenario	Assumption Analysis
Assume 50% cases, asymptomatic	The 'what if' analysis assumption of 50% asymptomatic, we will get 801,576 total cases, including half being asymptomatic. This will give us about 0.24% hospitalization as a % of population, in line with Influenza.
Assume 10 times more reported cases	The 'what if' assumption of 10 times more cases exist than reported, we will have only 1.03% hospitalized with critical symptoms. This is in line with the low-end estimate of influenza.
Assume 30 times more reported cases.	The 'what if' assumption of 30 times more cases exist than reported, the deaths as a % of sick will be at 0.11%, similar to the high-end estimate of influenza rates.

*Note: Death and hospitalization are considered as constants, since these cannot change, i.e. death and hospitalization cannot go unreported.

Here is a quick look at the above-mentioned calculations -

	Influenza Count (Oct to March)		COVID-19			
	Low Estimate	High Estimate	As of April 7	With 30% Asymptomatic	What if 10x	What if 30x
Illness	3,900,000	5,500,000	4,00,788	8,01,576	40,07,880	1,20,23,640
Medical Visits	1,800,000	2,600,000				
Hospitalization	400,000	7,30,000	41,000	41,000	41,000	41,000
Deaths	24,000	63,000	12,901	12,901	12,901	12,901
Illness as % of exposed Population	15.16%	21.37%	0.1212%	0.2425%	1.2125%	7.3880%
Medical Visits as % of exposed Population	7.00%	10.10%				
Hospitalization as % of exposed Population	0.16%	0.28%	0.0124%	0.2425%	1.2125%	7.3880%
Deaths as % of exposed Population	0.01%	0.02%	0.0039%	0.0039%	0.0039%	0.0039%
Hospitalization as % of sick	1.03%	1.33%	10.23%	5.11%	1.02%	0.34%
Deaths as % of sick	0.06%	0.11%	3.22%	1.61%	0.32%	0.11%

Additionally, if we look at the Europe data as a whole, as of April 8th, the total COVID-19 death rate as a % of population in the countries with more than 1000 reported cases is 0.0061%, that gives about 20,000 deaths total in the USA. Even if it's off by 3 times at 0.0183, which only 2 European countries (Italy and Spain at about 0.025%) have crossed, and may only cross, it bring the death count in USA from COVID-19 to 60,000, less than the already reported high end numbers for Influenza in USA for this season. Even if we look at USA at current Italy and Spain rates, the total death rate will be about 84,000.

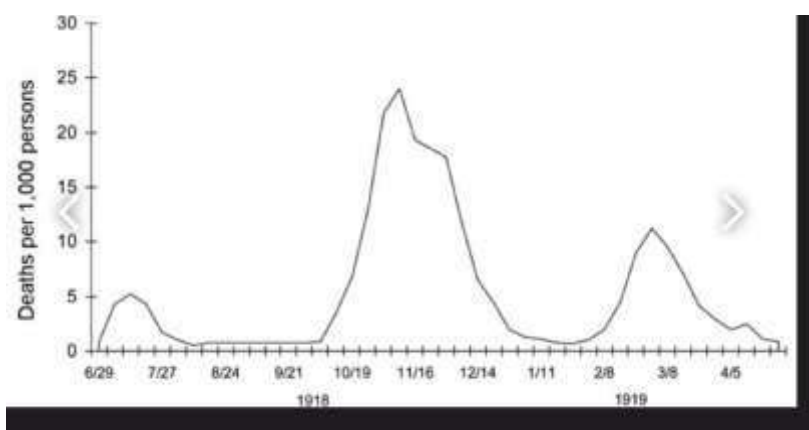
If we go by the worst-case predictions, then an epidemic seed on January 20th implies that by March end, about six million people in the US would have been infected. This is not the case. We were around 200,000 or below by March 31. Based on the current trends USA is more or less around the peak, so what would be the additional fatalities from the

current 12,901? Hundreds of thousands as claimed by some? Unlikely.

5.3 Comparison between Spanish Flu 1918 and COVID-19

The "Spanish" influenza pandemic of 1918–1919 had one third of the world's population (or ~500 million persons) were infected and had clinically apparent illnesses. The disease was exceptionally severe. Case-fatality rates were >2.5%, compared to <0.1% in other influenza pandemics. Total deaths were estimated at ~50 million and were arguably as high as 100 million.

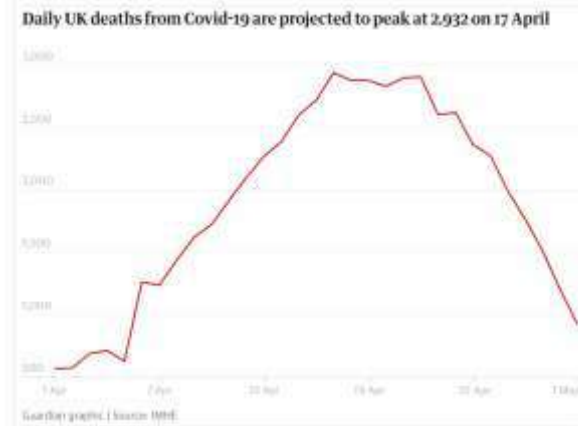
In one month alone, October 1918, there was 195,000 deaths by Spanish flu in USA. By today's population equivalent, death by COVID in that month would be 612,000. The total death so far in USA (over 100 days) by COVID is 22,106.



Here is the death per thousand charts for Spanish Flu for UK. This shows a high peak of about 25 deaths per thousand in a week. Highest death per thousand population per day recorded for COVID is 0.00093, across the globe, not just UK.

Comparing that to COVID-19 fatality chart from UK, if we use the peak week, we get an estimated death of 20,000 or 0.3 deaths per thousand. At peak, Spanish Flu was over 80 times more fatal in UK than peak of COVID.

*Used the daily chart – and it shows a peak of 29,000

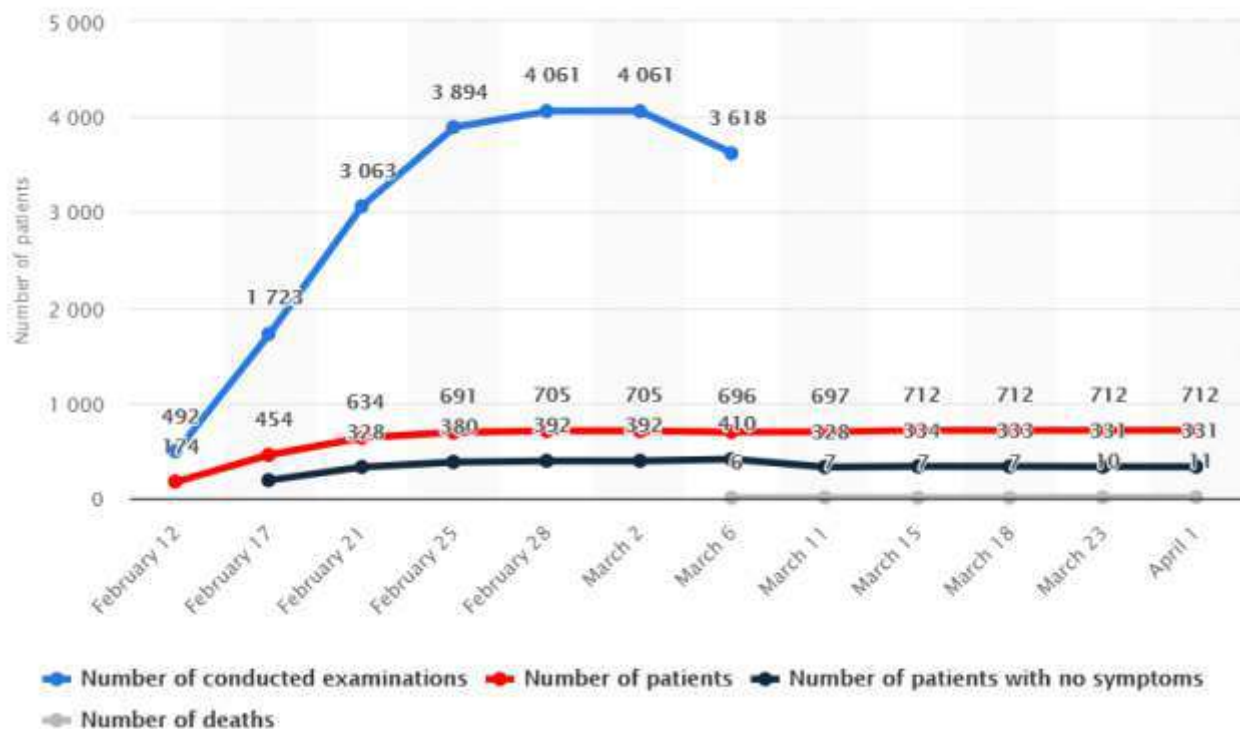


6. Case studies for major COVID-19 affected regions

6.1 Diamond Princess Cruise Ship and Hubei Province China

Analysis of Hubei province, where Wuhan is, from where COVID-19 is assumed to have started and Diamond Princess, a cruise ship which was isolated for weeks due to Coronavirus gives us early insights to infection rate and asymptomatic rates.

CDC reported that around 712 (19%) on of the total onboard Diamond Princess of 3711 were COVID-19 positive, of which 7 have died (0.98%). Of these 17.9% patients were asymptomatic.



Diamond Princess passengers 2666 in total had a median age of 69, with about 60% of the passengers over 60 years of age. The median age of crew was 36 years old. Analysis of the age demographics and passenger/crew ratio shows that only 10% of the infected cases were less than 40 years in age Hubei province, from a total of 5,72,37,740 population, reported 67,802 confirmed cases (0.12%) and 3,203 deaths (4.72%).

From a statistical perspective, there are questions about the accuracy of Hubei province data, however, Diamond princess data has been validated.

6.2 Europe





































As of April 6, there are 26 countries with over 1000 COVID-19 cases reported. Italy and Spain have

the maximum number of deaths. Europe data is quite broad, and includes the currently worst impacted Italy, Spain, France & UK etc., broad geographic terrain, many different geographical, economic and social backgrounds and response action. (example Sweden has not imposed any restrictions on citizens so far)

WHO on March 31 has reported that they expect Europe to be at or crossing the peak point.

The overall sickness ratio of Europe as a percentage of total population is 0.091%. Spain and Italy have these rates more than 0.25%. If the top 6 countries are analysed together separately versus the rest of others together, there are clear distinctions. The death ratio of the top 6 have over 10x more than the rest of 38.

Country Groups	Total Population	COVID-19 Confirmed	COVID-19 Deaths	Sick Ratio	Death Ratio of Population	Death Ratio of Sick
Top 6 impacted	26,91,00,621	4,56,682	52,305	0.1697%	0.0194%	11.4533%
Bottom 38 impacted	47,81,93,955	2,19,694	5,293	0.0459%	0.0011%	2.4093%

	Country	Population	Confirmed Sick	Deaths	Sick Ratio of Population	Death Ratio of Population	Death Ratio of Sick
	Italy	6,04,61,826	1,35,586	17,127	0.224%	0.028%	12.632%
	Spain	4,67,54,778	1,46,690	14,555	0.314%	0.031%	9.922%
	France	6,52,73,511	78,167	10,328	0.120%	0.016%	13.213%
	United Kingdom	67886011	55,242	6,159	0.081%	0.009%	11.149%
	The Netherlands	17134872	18,803	2,101	0.110%	0.012%	11.174%
	Belgium	1,15,89,623	22,194	2,035	0.191%	0.018%	9.169%
	Germany	8,37,83,942	1,07,659	2,017	0.128%	0.002%	1.874%
	Switzerland	86,54,622	21,657	765	0.250%	0.009%	3.532%
	Sweden	10099265	7,693	591	0.076%	0.006%	7.682%
	Portugal	10196709	11,730	311	0.115%	0.003%	2.651%
	Austria	90,06,398	12,206	220	0.136%	0.002%	1.802%
	Denmark	57,92,202	4,681	187	0.081%	0.003%	3.995%
	Ireland	4937786	5,364	174	0.109%	0.004%	3.244%
	Romania	19237691	4,057	157	0.021%	0.001%	3.870%
	Poland	37846611	4,413	107	0.012%	0.000%	2.425%
	Greece	10423054	1,832	81	0.018%	0.001%	4.421%
	Czech Republic	10708981	4,735	78	0.044%	0.001%	1.647%
	Norway	5421241	5,768	76	0.106%	0.001%	1.318%
	Serbia	8737371	2,200	58	0.025%	0.001%	2.636%
	Hungary	9660351	895	50	0.009%	0.001%	5.587%
	Russia	145934462	6,343	47	0.004%	0.000%	0.741%
	Ukraine	43733762	1,319	38	0.003%	0.000%	2.881%
	Luxembourg	625978	2,804	36	0.448%	0.006%	1.284%
	San Marino	33931	277	32	0.816%	0.094%	11.552%
	Slovenia	2078938	1,021	30	0.049%	0.001%	2.938%
	Bosnia and Herzegovina	3280819	674	29	0.021%	0.001%	4.303%
	Finland	5540720	2,176	27	0.039%	0.000%	1.241%
	North Macedonia	2083374	599	23	0.029%	0.001%	3.840%
	Bulgaria	6948445	581	23	0.008%	0.000%	3.959%
	Andorra	77265	526	21	0.681%	0.027%	3.992%
	Albania	2877797	377	21	0.013%	0.001%	5.570%
	Estonia	1326535	1,108	19	0.084%	0.001%	1.715%
	Moldova	4033963	965	19	0.024%	0.000%	1.969%
	Croatia	4105267	1,222	16	0.030%	0.000%	1.309%
	Lithuania	2722289	843	14	0.031%	0.001%	1.661%
	Belarus	9449323	700	13	0.007%	0.000%	1.857%
	Iceland	341243	1,562	6	0.458%	0.002%	0.384%
	Slovakia	5459642	534	2	0.010%	0.000%	0.375%
	Montenegro	628066	233	2	0.037%	0.000%	0.858%
	Latvia	1886198	542	1	0.029%	0.000%	0.185%
	Liechtenstein	38128	77	1	0.202%	0.003%	1.299%
	Monaco	39242	73	1	0.186%	0.003%	1.370%
	Malta	441543	241	0	0.055%	0.000%	0.000%
	Vatican City	801	7	0	0.874%	0.000%	0.000%
	Total	74,72,94,576	6,76,376	57,598	0.091%	0.008%	8.516%

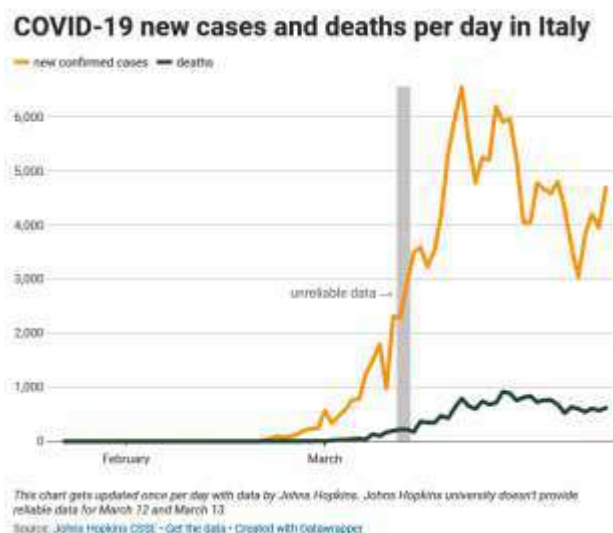
Summary table of confirmed cases in Europe (as of 6 April 2020)

6.3 Italy Detailed Analysis

Detailed analysis shows that certain areas of Italy was worse affected than others. The peak number of cases could not be managed by the available and healthcare capacity, causing large rate of deaths in the early stage.

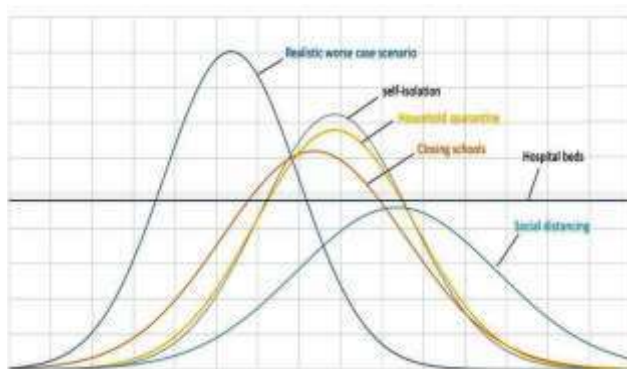
Region/Province	Population	Cases		Deaths		
		#	% of Population	#	% of Population	% of cases
Lombardy	1,00,60,574	52,325	0.520%	9,484	0.094%	18.13%
Emilia-Romagna	44,59,477	17,825	0.400%	2,180	0.049%	12.23%
Piedmont	43,56,406	13,343	0.306%	1,319	0.030%	9.89%
Veneto	49,05,854	11,925	0.243%	695	0.014%	5.83%
Tuscany	37,29,641	6,173	0.166%	369	0.010%	5.98%
Liguria	15,50,640	4,757	0.307%	620	0.040%	13.03%
Marche	15,25,271	4,710	0.309%	630	0.041%	13.38%
Lazio	58,79,082	4,149	0.071%	238	0.004%	5.74%
Campania	58,01,692	3,148	0.054%	216	0.004%	6.86%
Apulia	40,29,053	2,514	0.062%	209	0.005%	8.31%
Trentino - South Tyrol	10,72,276	4,287	0.400%	418	0.039%	9.75%
Friuli-Venezia Giulia	12,15,220	2,153	0.177%	164	0.013%	7.62%
Sicily	49,99,891	2,097	0.042%	125	0.003%	5.96%
Abruzzo	13,11,580	1,799	0.137%	172	0.013%	9.56%
Umbria	8,82,015	1,263	0.143%	49	0.006%	3.88%
Sardinia	16,39,591	935	0.057%	52	0.003%	5.56%
Aosta Valley	1,25,666	835	0.664%	100	0.080%	11.98%
Calabria	19,47,131	833	0.043%	60	0.003%	7.20%
Basilicata	5,62,869	291	0.052%	14	0.002%	4.81%
Molise	3,05,617	224	0.073%	13	0.004%	5.80%
Italy	6,03,59,546	1,35,586	0.225%	17,127	0.028%	12.63%

Italy is assumed to have crossed the peak on March 28th



assuming others may be sick, keep about 6 feet away from them and use a mask. Wash hands often with soap and water for at least 20 seconds especially after been in a public place, or after blowing nose, coughing, or sneezing. If soap and water are not readily available, use a hand sanitizer that contains at least 60% alcohol. Cover all surfaces of hands and rub them together until they feel dry. Also, avoid touching eyes, nose, and mouth with unwashed hands.

Some models have shown that masks and social distancing together can reduce the spread rate by as much as 50%. Combined this with available Healthcare capacity, most regions can manage the spread effectively.



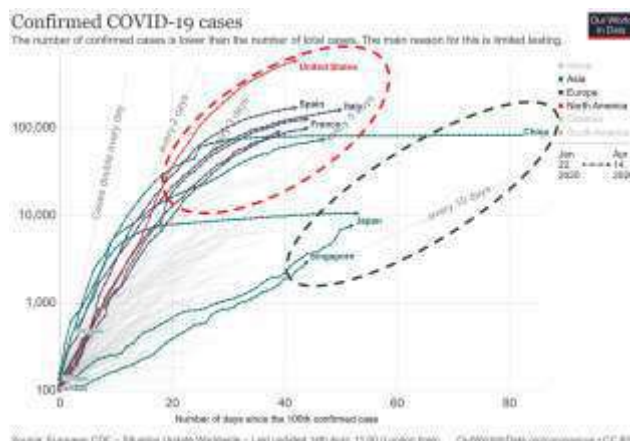
Masks don't just offer protection to the wearers. They protect everyone else around them by controlling the source of the virus, which can travel as far as 2–6 meters when an infected person coughs or sneezes. Even a basic face-covering like a scarf provides some protection for everyone else in the vicinity.

If we analyse cultures which practice physical distance and use masks when sick Japan, China, South Korea and Singapore we find that the rate of progression there is much slower. Japan as of date does not have any major controls, with public transport operating as normal. In contrast if we analyse Italy, Spain, Germany USA, UK, Iran, 20

we find that the rate of reproduction there is much faster. In short, cultures that use mask have 'flattened the curve' much better.

All wearing masks and keeping social distance

can only be diligently followed by citizens in a highly disciplined society. Otherwise, the governments will have to bring in draconian conditions like lockdown.



7.3 Herd Immunity

If we can test all the population and isolate those with the disease as soon as possible, we can reduce the spread of the disease. However, lack of availability of test kits make this impossible under current circumstances. However, this won't prevent the risk of infection to the susceptible population and stop transmission. For COVID-19 to end its transmission, R_0 has to reduce by increasing the immunity of the broader set of population. This reduces the susceptible population and stops the spread. This is known as Herd immunity or the resistance to the spread of a contagious disease within a population that results if a sufficiently high proportion of individuals are immune to the disease, especially through vaccination.

As vaccines are not available, the only way to get immunity is for the human body to have resistance through exposure of the disease. With an R_0 of 2.2, if 55% of population is immune, then the disease spread will stop. Also, are additional factors that bring immunity like age, gender, blood group, genetics and health conditions. Resistance to other illnesses like malaria, tuberculosis, sufficiency of Vitamin D etc may also bring in additional immunity, although there is not enough studies to substantiate these. Also, there may be broader aspects like temperature, humidity etc., that can reduce

coronavirus spread. So, the chances are that a large percentage of population will already have immunity. As fewer people get sick, disease will stop its spread and with proper controls, the disease will die down.²¹

7.4 Lock down and other draconian controls

Lockdown does not eliminate the epidemic but will only reduce the rate of transmission of the disease. The number of people at risk (susceptible population) for the disease is not reduced through lockdown. As the immunity does not increase, and the disease is continuing actively in other states and around the world, we cannot eliminate the risk of the disease, regardless of how long the lockdown lasts.

Lockdown does not eliminate coronavirus threat. Due to low community contact during the lockdown, the rate of outbreak rate will be reduced.

Lockdown does have its benefits. With the principle of "Flatten the Curve", it can reduce the number of infected people at the same time and ensure that the health system is adequate to handle them. Lockdown does not reduce the total number of patients, but postpone the disease and prolong the duration of the disease in the society. Therefore, ideally, instead of all susceptible citizens getting sick together, burdening and breaking the system, if they are spread over a time, and peak count is maintained below the system capacity, the sick can be taken care of in the healthcare system. Flattening the curve does not reduce the probability count of the total number of sick, it only spreads the disease over a longer period. As the peak reduces, the active epidemic period increases proportionally.

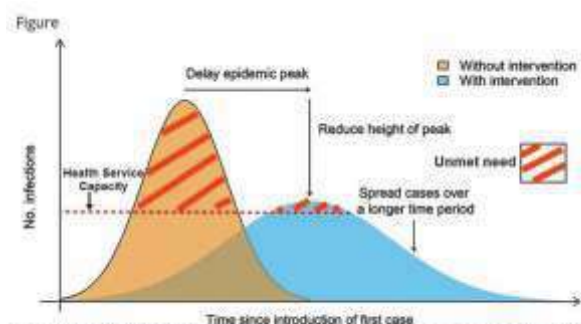


Figure 1: Intended impact of enhanced hygiene and social distancing measures on the COVID-19 pandemic adapted from Fong.⁶

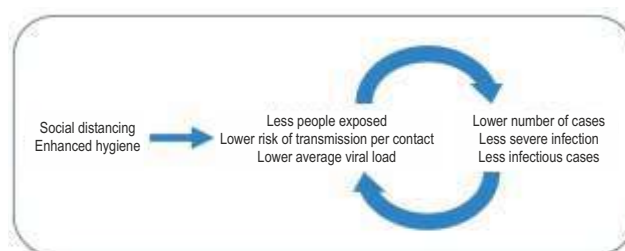


Figure 2: Conceptual model of how pre-emptive interventions with a negative multiplier effect could impact an impending epidemic

However, if the peak is much lower than the capacity of the health sector, the healthcare system will not be utilized properly. This is not ideal either. As only a few people got sick, the immunity level has not increased, the susceptible population is more or less same as January 1st. If the illness returns after the lockdown, almost all population is at risk again. Also, the health care capacity is not fully utilized.

In summary, if the curve is flatted too much, the risk has not reduced, just delayed (As Kicking the Can Down the Road). If the disease recurs, we will have to manage the same susceptible population we are managing now. If the healthcare capacity does not increase during this period, peak risk will remain the same.

Here is the key difference with Nipah, as Nipah was a controlled set of infection from a single index patient and once that disease spread is eliminated through isolation, there are no chances of subsequent waves.

7.5 Optimum Flattening of the Curve

Fattening the curve was needed in Wuhan, Spain, Italy, etc., as these areas did not have enough Healthcare capacity. Lombardy in Italy, Madrid in Spain and Wuhan in China are the most badly affected parts of the world so far. For example, in Italy, during the peak on March 28th on the need for ICU beds for critical patient care was 6,562, but the available ICU beds was only 2,059. This shows severe under capacity, which must be flattened to reduce the death rate. Through flattening of the curve, the volume of sick and needing critical care was reduced, keeping under the healthcare system capacity and ensure available services to citizens.



However, in certain other countries like Japan, physical distancing and other controls have ensured that the curve remained flat and government has not taken draconian measures.

7.6 The threat of over flattening the curve

Here's an example to show how "Flatten the Curve" is achieved through lockdown and how it does not eliminate the underlying problem. Imagine we are having a wedding party for a 1000 people. But if the dining hall can only accommodate 100 people, we have a 'resource constrained' situation. This is not good as it can cause chaos. To overcome this, we get 100 at a time and, will serve ten rounds. If each round takes an hour, 100 people can be served 10 times and the entire 1000 can have covered in 10 hours. Just that we only accommodate 100 instead of 23.

1000 per round and it takes 10 hours in total versus 1 hour. We use this kind of "Flatten the Curve" technique in many situations in our daily lives to optimally utilize resources.

But the trouble is, requesting people to voluntarily wait for meal might not work practically. They will all want to have food in the first round itself and crowd the dining hall and create disaster. In that case, the host could ask everyone who comes to the wedding, stay locked in the wedding hall, to allow in 100 people at a time to Dining hall. Dining will

complete in batches of 100 people in 10 hours. But what if the over restrict and allow only 10 people into each batch in dining hall? Instead of batches of 100, we may have batches of 10 and the whole process takes 100 hours. While people are waiting hungry, we have wasted seats. I call it as "over-flattening the curve". This situation is not good either. Here we are in a state of 'resource underutilization' and completion time is extended. These are just the things we do in every day in business management.

An R0 factor of 2.2 suggests that we will only reach herd immunity to Covid-19 when approximately 55% of the population is immune - remember that immunity is currently only reached by getting the infection as we have no vaccine! Naturally many people are coronavirus-resistant for many reasons. Our inherent immunity, age, gender, blood group, genetics, resistance to other illnesses, health status, pre-vaccination history and so on. Then there may be the temperature, humidity, and such factors might impact the spread. Then a good percentage of people will already have immunity.

Returning to our example, many of the 1000 people don't need food as some ate at home, some brought food, some does not eat from the outside, etc. If we can get a count of only those who will eat, then it is easy to manage the dining hall. Similarly, if we could, we would all be able to detect people who

are at risk of getting infected by a test and thereby our susceptible population. Not the whole population. Just as of the 1000, only a portion are hungry and will have the meal. What if, of the entire 1000 attendees only 100 or fewer are going to eat? In that case, we can sit down the entire guests and complete the entire meal in an hour. But we don't have the ability to test so many of our population, so we have to make some models and assumptions. Just as when we invite for wedding, we make an assumption of how many might attend from each family, how many might eat, how many adults, children, vegetarian non-vegetarian etc. Similarly, we have to look at the data from other places where

they have this disease, and what they have experienced.

Coronavirus started, in Hubei province of China. City of Wuhan was the earliest impacted. However, data from China is assumed to be not fully reliable hence not used for modelling. Outside of China, Italy is one of the earlier impacted countries, and the most affected by coronavirus in Europe. It is assumed that the peak of impact passed Italy on March 27th. So far, as of April 07, 1,51,226 people have been infected. It is assumed that by the end of this cycle, by August 4, about 170,000 people will be impacted. That is about 0.27% of the population of Italy.

Italy Province Groups	Population	% of Population	Cases		Deaths		
			#	% of Population	#	% of Population	% of cases
Top 7 impacted	3,05,87,863	50.68%	1,11,058	0.36%	15,297	0.050%	13.77%
Bottom 13 impacted	2,97,71,683	49.32%	40,168	0.13%	1,830	0.006%	4.56%
TOTAL	6,03,59,546	100.00%	1,51,226	0.25%	17,127	0.028%	11.33%

Remember, Italy is the most affected area in the world, probably due to the age of the people. With the exception of the 6 most affected countries in Europe including Spain, Italy, France and UK,

the infection rate of the other 38 countries is only about a quarter. And less than a fifth of the rate of Italy (as of April 7)

Europe Country Groups	Total Population	Population %	COVID-19 Confirmed	COVID-19 Deaths	Sick Ratio	Death Ratio of Population	Death Ratio of Sick
Top 6 impacted	26,91,00,621	36.01%	4,56,682	52,305	0.1697%	0.0194%	11.4533%
Bottom 38 impacted	47,81,93,955	63.99%	2,19,694	5,293	0.0459%	0.0011%	2.4093%
TOTAL	74,72,94,576	100.00%	6,76,376	57,598	0.0905%	0.0077%	8.5157%

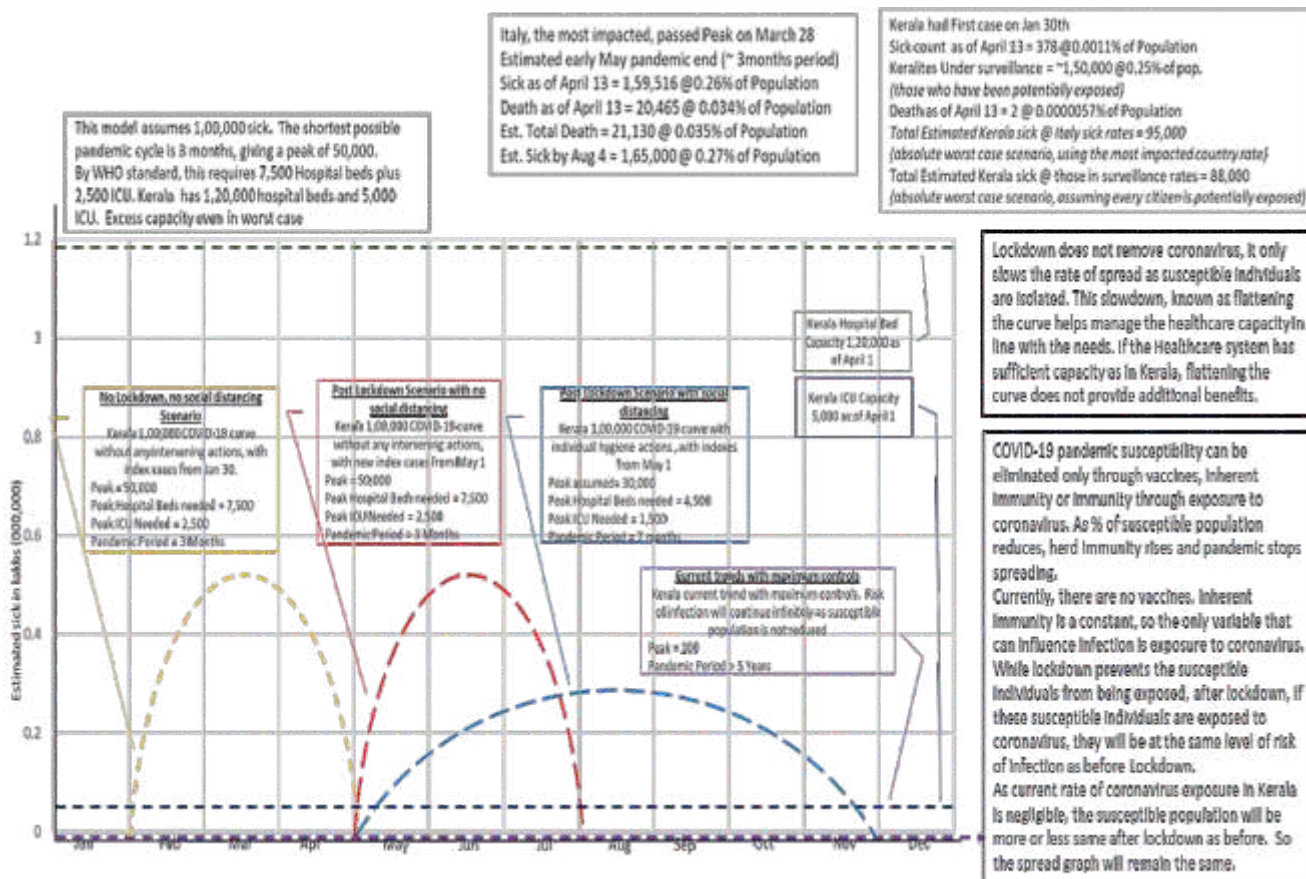
According to the WHO, 80% of COVID-19 sufferers have only a small amount of difficulty. 15% reported inpatient care and 5% required intensive care. What do these figures mean?

Kerala State Minister of Health informed that in early April that there are 1.2 lakh hospital beds and 5,000 critical care beds available in Kerala for COVID-19 emergencies. 1.2 lakhs beds can cover 8 lakh patients at a time, assuming the WHO hospital rate of 15%, 8 lakhs at peak means a total of 16 lakh or more patients can be covered in a 3 month period. If we look at ICU, 5,000 ICU beds can cover 1 lakh patients one time or 2 lakh patients over a 3 month period. This is Kerala's patient handling capacity as of April 1.

Assuming the same severe infection rate as Italy, another place with 3,50,00,000 population, like Kerala, less than 1,00,000 people will get sick in a

worst case scenario. A total of 15,000 hospital beds and 5,000 intensive care facilities will be required. If assumed at the rate of other countries in Europe, the total sick would be 20,000 patients, needing 3,000 hospital beds and 1000 intensive care beds. But many experts say that India's immunity in general is much higher than in Europe. In that case we are looking at lesser numbers.

Further deep diving Kerala numbers. As of April 13, Kerala with a population of about 3,50,00,000 had cumulatively about 1,50,000 or more in isolation – current active is 1,12,83. (as health department do not publish cumulative number of those kept in surveillance, this number was derived from daily trends and the tele counselling number of 1,95,161). Meaning, they had been directly exposed one way or other to an environment of coronavirus -air traveller, contact etc. Furthermore, 15,683 test results came back, of which only 378 was positive.





The sick count of 378 is 0.0011% of active population or 0.25% of those under surveillance – or exposed population. However, we can argue that lockdown and other controls have shielded vast majority from getting exposed one form or other, let us use the under surveillance as they would have been more or less exposed. Using this logic, even if we expose all of Kerala to COVID-19, the potential infection would be 87,267 or about 13,000 hospital admissions and 4,360 will require critical care by WHO standards. These worst-case scenario numbers are below Healthcare available capacity and hence we can provide necessary and sufficient care to those infected.

With only 378 patients as of April 13th, by not exposing any meaningful percentage of population to virus and thereby immunity, are we not "over-flattening the curve" and pushing the pandemic risk

to a future date? 378 patients show that Kerala is pretty much remaining at the same level of pandemic risk as January 1.

The disease pattern shows an approximate 3-month wave pattern during the statistically significant period. If we round up the potential sick to the nearest 00,000, we get 1,00,000 sick and in a 3 month period, the peak will be 50,000. By WHO standard, the peak will require, 7,500 hospital beds and 2,500 ICU Beds.


Many places in the world are confronting this disease in different ways. Sweden in Europe has no restrictions, for example. Meanwhile Belgium, which is roughly the same population, has implemented a very rigorous lockdown. Looking at the current projection models, there is no difference between these countries from COVID-19 when the season is over.

Country	IHME Model Prediction			Government Controls			
	Estimated Peak	Estimated death by August 14	Estimated % Death	Stay at Home	School Closure	Non Essentials Closure	Emergency Travel only
 Belgium	April 6, 2020	3546	0.031%	March 18, 2020	March 14, 2020	March 18, 2020	March 18, 2020
 Sweden	April 25, 2020	4182	0.041%	NO	NO	NO	NO

Data Projection as of April 10 based on University of Washington IHME

In the United States, Arkansas and Mississippi are neighbouring similarly populated states. Arkansas implemented school closures and no other major control, while Mississippi implemented a very

stringent lockdown. Looking at the current projection models, there is no difference between them in the impact from COVID-19 when the season is over.

USA State	Population	Sickness			Death		
		First Case	Confirmed Sick	Sick Ratio of Population	Deaths	Death Ratio of Population	Death Ratio of Sick
 Mississippi	29,76,149	March 11, 2020	2,260	0.076%	76	0.003%	3.363%
Arkansas	30,17,825	March 11, 2020	1,146	0.038%	21	0.001%	1.832%

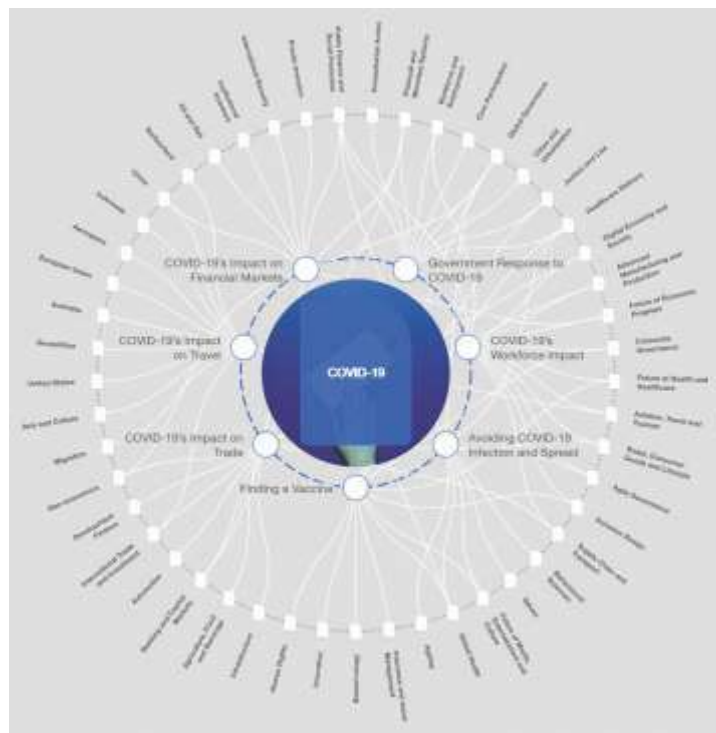
Official reports from state health official records- Data as of April 10

In Japan, schools were closed and major events where large crowd gather is stopped. But in most parts of the country, people's lives are more or less normal. However, people make good use of the mask. Only 5% (1/20) of the disease rate of in Europe is observed in Japan.

Hence it can be argued that Lockdown and similar draconian measures are not necessary if the Healthcare system has adequate capacity and if the citizens establish a level of voluntary control incusing wearing masks and keeping social distance.

The impact of lockdown may cause more deaths than the coronavirus, especially in developing countries. In addition, lockdown will impact livelihood, lead to a financial crisis and a long-term hunger. Good public policy is always based on the best science, the latest and most reliable evidence. For risk management, we need to look at probability and impact. While the promulgated impact and probability of COVID-19 is catastrophic, realistic look at numbers show that it's not that damaging.

8. Impact of COVID-19 – Beyond Health



World Economic Forum and other credible organizations have captured the potential impact of COVID-19. This includes Healthcare related as well as other areas including Government Response, Trade, Workforce, Geopolitics, Financial Markets, etc. This document tries to capture Economic impact and Social Impact, in context of Kerala.

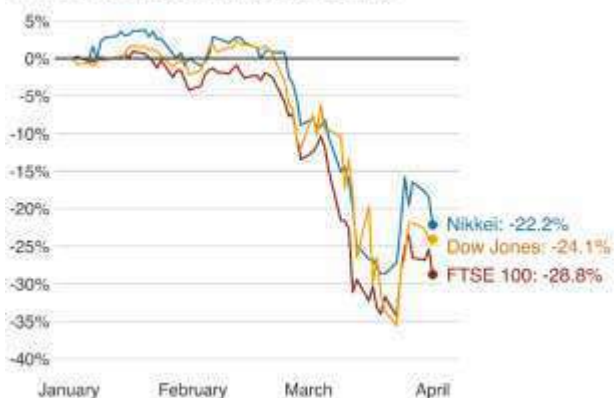
8.1 Overall economic analysis

Our primary focus now must be to support the health sector, quickly increasing the capacity of

beds, staff and ventilators, mobilising the production of the valuable equipment through tax breaks, public purchases or direct government control, and spurring innovation in treatments and vaccines.

However, the economic impact must not be ignored. Coronavirus will bring a major market shock as it will impact both demand and supply. As the global markets are already impacting from low oil prices, this causes major resonance to market causing recession in many countries.

The impact of coronavirus on stock markets since the start of the outbreak



Source: Bloomberg, 01 April 2020, 09:00 GMT

Acuite Ratings & Research earlier this month estimated that the lockdown will cost the Indian economy almost \$4.64 billion (over Rs 35,000 crore) every day and the entire 21-day lockdown will result in a GDP loss of almost \$98 billion (about Rs 7.5 trillion

World Bank has said that India is likely to record its worst growth performance since the 1991 liberalisation this fiscal year as the coronavirus outbreak severely disrupts the economy. India's economy is expected to grow only 1.5% to 2.8% in the 2020-21 fiscal which started on April 1, the World Bank said in its South Asia Economic Focus report. It had estimated India will grow 4.8% to 5% in the 2019-20 fiscal that ended on March 31.

It also predicts the South Asian region, comprising eight countries, growing by only 1.8 - 2.8% this year, down from the 6.3% it projected six months ago.

The Asian Development Bank (ADB) sees India's economic growth slipping to 4% in the current fiscal, while S&P Global Ratings has further slashed its GDP growth forecast for the country to 3.5% from a previous downgrade of 5.2%.

Fitch Ratings puts its estimate for India growth at 2%, while India Ratings & Research has revised its FY21 forecast to 3.6% from 5.5% earlier.

International Monetary Fund (IMF) has warned of a recession worse than the global financial crisis of 2008, when the e global economy lost more than \$2

trillion in value, a drop of nearly 4%. This will impact manufacturing and Services sector, including tourism and IT, and can cause massive job layoffs including abroad.

Moody's downward revised India's GDP growth for 2020 calendar year to 2.5% from 5.3%.

Barclays estimates that India's aggressive 21-day lockdown could bring the country's growth down to 2.5% from the 4.5% it had earlier estimated.

Poor economic conditions impact consumer confidence and consumer spending. India's Private Consumption accounted for 62.4 % of its Nominal GDP in Dec 2019.

8.2 Impact to MSMEs, Service & Unorganized Sectors

Services sector is the largest sector of India. Gross Value Added (GVA) at current prices for Services sector is estimated at 92.26 lakh crore INR in 2018-19. Services sector accounts for 54.40% of total India's GVA of 169.61 lakh crore Indian rupees. With GVA of Rs. 50.43 lakh crore, Industry sector contributes 29.73%. While, Agriculture and allied sector shares 15.87%.

MSMEs are likely to be the biggest losers. These companies are too small to have enough of a cushion to last through a pandemic like this one and cannot survive a prolonged shutdown. MSME sector contributes 29% to India's GDP and provides employment to more than 120 million persons and contributed 48.1% to India's total exports in 2018-19.

How Americans are changing their plans during the coronavirus pandemic

	March 13-15	March 16-19	March 20-22
Avoided going to events with large crowds, such as concerts, festivals, or sporting events	59%	79%	92%
Avoided travelling by airplane, bus, subway, or train	55%	75%	87%
Canceled or postponed travel plans	39%	57%	73%
Avoided going to public places, such as stores or restaurants	30%	54%	72%
Avoided small gatherings of people, such as with family or friends	23%	46%	68%
Stocked up on food, medical supplies, or cleaning supplies	39%	52%	59%

Source: Gallup Panel, 2020

Vox

Kerala has the 12th largest number of MSMEs in India - around 2.4 million units generating employment for approximately 4.5 million people.

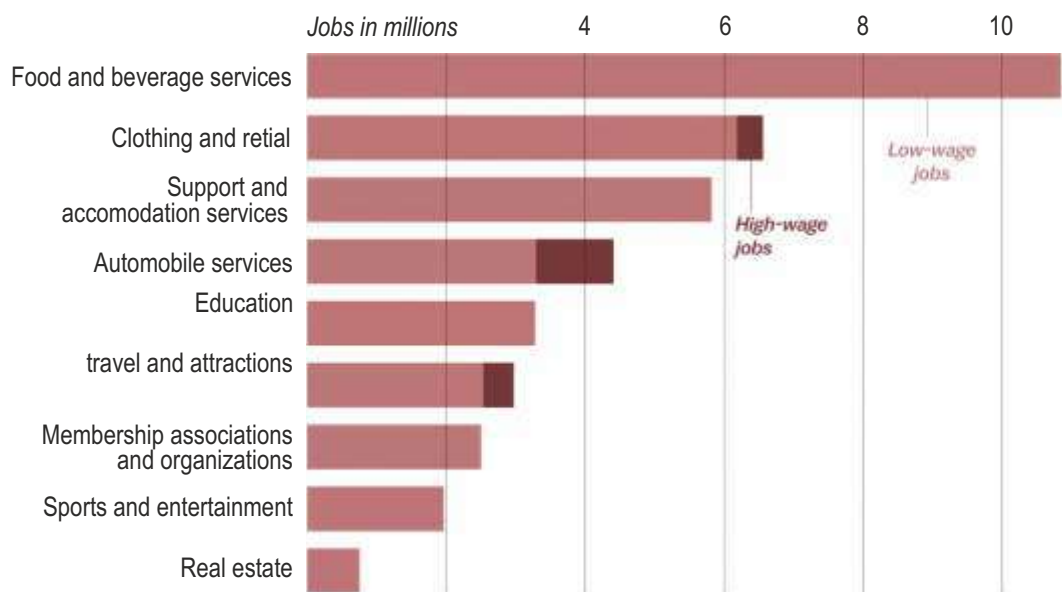
Kerala has about 44% of its labour force as casual workers. Also, about 3.5 million migrant laborers are operating as casual workers. 0.7 million workers are factory employees. Daily wage earners will be hurt badly.

3.1 per cent of Kerala's population work in Fishing

and allied sectors. They reside in 222 marine fishing villages and 113 inland fishing villages of the State. The fish worker population of the State in 2016-17 is estimated to 1.03 million

Around 3 million Keralites are working abroad, mainly in middle east and a large percentage of them belong to lower-income or middle income group. Remittances contribute nearly 36.5% to the state net domestic product.

Jobs vulnerable to layoffs during the Covid-19 pandemic in the US by industry



Note: "Low income" jobs are defined as making a weekly income of less than \$801.47. Values are rounded.
Source: US Private Sector Job Quality Index

Vox

The impending recession and oil price slump will impact many of these jobs, thereby remittances and moreover, most of these non-residents will move back to Kerala post-coronavirus. This will adversely impact the economic condition of Kerala both directly and indirectly and will affect the social.

Agriculture and allied sectors contributed 10.5% of in total GSVA of Kerala in 2016-17

All these sectors will be adversely impacted

Added challenge to India is the impact of China. They've already began production. As of March 25, China seems to be getting back to business, with

important indicators like electricity and steel demand and auto manufacturing returning to levels not far from normal. China's head start could be devastating blow to many Indian enterprises.

9. Social Impact of Lockdown

The WHO has asked governments to test, trace, and isolate, as well as scale up healthcare system capacity and capability for care. However, governments have opted for a national lockdown to compensate for the lack of adequate test and healthcare system capacity.

Lockdown measures will delay pandemic spread

and may save many lives in places where through 'flattening of the curve' healthcare capacity can be managed in line with the need for hospital care. Improved air quality and a reduction in traffic accidents will save many more lives as well. But the pandemic will have a severe impact on the economy over a number of years. Centre for Global Development has stated that the impact of lockdown of India, if fully felt, will kill far more from indirect consequences than from COVID 19.

Anxiety, frustration, panic attacks, loss or sudden increase of appetite, insomnia, depression, mood swings, delusions, fear, and suicidal tendencies, have become quite common during the lockdown.

According to India's National Commission for Women, there has been a sharp rise in the number of distress calls during the coronavirus lockdown. As per reports from New Jersey USA, domestic violence has risen up to 25% during lockdown.

Social scientists assume that recession and financial loss can double the suicide rate.

There have been many reported cases of suicide due to alcohol withdrawal symptoms.

An estimate from United States have that if the jobless rate rises to 20%, it could translate into 48 million years of lost human life through various reasons.

Additionally, due to transportation issues, lack of medicines and supplies, lack of access to healthcare many lives have already been lost during lockdown

10. Conclusion

Analysis of Data so far shows us that

- ▶ COVID-19 is a very dangerous pandemic; however, the insinuated death rates are not substantiated by the experience in many countries that have gone through the peak phase. Current predictions have the death toll of COVID-19 in USA to be more or less in line with seasonal Flu, or influenza.
- ▶ The disease has crossed or is its peak its peak in China, Southeast Asia, most of Europe, parts of

USA, etc. So, large portion of world population has already crossed through the worst phase of this disease.

- ▶ Not all COVID-19 patients require hospital treatment. WHO predicts that 80% of those infected does not need any hospital care, 15% may require normal care at hospital and 5% might require critical care.
- ▶ WHO has confirmed that COVID-19 can only be prevented through immunity that comes from vaccination, inherent immunity or exposure to disease. Vaccination is not currently available and might a few years. So, the susceptible population will not come down without exposure to disease.
- ▶ Coronavirus spreads from droplets. Wearing mask and keeping physical distance, washing hands can help control the spread. Social distancing and use of masks are sufficient in most cases to manage the spread.
- ▶ Once a patient is identified, effective management including quarantine, treatments as per protocols, contact tracing to identify exposed individuals and subsequent actions can be effective in disease spread control.
- ▶ Lockdown will not prevent COVID-19. It will only shield the susceptible population during the lockdown period and delay the spread. Once lockdown is over, the spread will continue through the susceptible population, in case of any exposure.
- ▶ Comparison of similar locations with draconian lockdown and no-control, or citizen restraints has shown that lockdown may not provide additional benefit other than flattening the curve for places with constrained healthcare capacity.
- ▶ 'Flattening the curve' through lockdown is not required in every country/region. Flattening the curve is only helpful when healthcare capacity is lower than the total number of infected, that require hospitalization, or 20% of the infected. Moreover 'over flattening' the curve causes longer period of potential outbreak period and

underutilization of healthcare capacity.

- ▶ As the infected rate in the countries that crossed peak is only a very small percentage of the population, eg. 0.27% of population in Italy, the worst impacted is infected. These are the carriers. Certain portions of Italy, Hubei Province, New York, Spain, UK etc., required lockdown to flatten the curve. Once the capacity is in line, lockdown provides no additional benefit. This means in a population of 3,50,00,000 only 378 are infected. 3,49,99,622 citizens are locked down for no reason. This will cause social issues, including fatalities.
- ▶ Extending lockdown does not ensure disease prevention. As COVID-19 is not an indigenous disease, nor is from controlled set of index patients, after lockdown is lifted, the spread will continue as susceptible population has not reduced.
- ▶ Lockdown causes unnecessary burden to the exchequer and loss of revenue. Free ration, free

food, free hospital care, ambulances, additional procurement all causes additional cost. Loss of sales tax income tax, excise revenue, lottery revenue all impact, all revenue adversely.

- ▶ The financial impact and the livelihood impact from coronavirus will be unprecedented. Lockdown will cause additional stress to the economy and will accelerate recession, inflation, bankruptcies, workforce reduction and income loss.

In short, there is minimal healthcare risk and potential substantial economic and social benefits of lifting Lockdown in Kerala. The recommendation is to lift lockdown in aggressive, yet progressive manner, encourage citizens to voluntary use masks, keep physical distance and disinfect hands frequently. Government must establish a war room to monitor the active and new cases, contact tracing, ensuring protocols and manage within Healthcare capacity.

Few Reference Links

- <https://www.who.int/>
- <https://www.cdc.gov/>
- <https://www.ecdc.europa.eu/en>
- [s://ourworldindata.org/](https://ourworldindata.org/)
- <https://www.ncbi.nlm.nih.gov/>
- <https://www.sciencemediacentre.org/>
- <https://www.businessinsider.com/>
- <https://www.vox.com/>
- <http://www.healthdata.org/>
- <https://www.forbes.com/>
- <https://www.wsj.com/>
- <https://www.nikkei.com/>
- <https://openknowledge.worldbank.org/>
- <https://www.cgdev.org/>
- <https://icmr.nic.in/>
- <https://www.reuters.com/>
- <https://www.weforum.org/>
- <https://theconversation.com/>
- <https://www.livescience.com/>
- <http://gabgoh.github.io/COVID/index.html>
- <https://www.thehindu.com/>
- <https://www.researchgate.net/>
- <https://www.theguardian.com/>
- <https://en.as.com/en/>
- <https://www.statista.com/>
- <https://dhs.kerala.gov.in/>
- <https://www.whitehouse.gov/>
- <https://www.cabdirect.org/>
- <https://www.jstor.org/>
- <https://www.researchgate.net/>
- <https://www.worldometers.info/>
- <https://www.health.com/>
- <https://www.helsinkitimes.fi/>
- <https://web.stanford.edu/>
- <https://www.jhu.edu/>
- <https://theconversation.com/>
- <https://cnn.com/>
- <https://worldpopulationreview.com/>
- <https://www.healthy.arkansas.gov/>
- <https://msdh.ms.gov/>
- <https://www.newyorker.com/>
- <https://www.usnews.com/>
- <https://www.euronews.com/>
- <https://sverigesradio.se/>
- <https://www.schengenvisa.info.com/>
- <https://www.theguardian.com/>
- <https://www.arcgis.com/>
- <https://who.sprinklr.com/>
- <https://www.hoover.org/>
- <https://www.medicaleconomics.com/>
- <https://www.dw.com/en/>
- <https://www.cebm.net/>
- <https://abcnews.go.com/>
- <https://www.statnews.com/>
- <https://www.aljazeera.com/>
- <https://www.barrons.com/>
- <https://www.census.gov/>
- <https://www.bloomberg.com/>
- <https://medium.com/>

New Normal and influence of Information Technology

Shri. Santhosh Kurup, CEO, ICT Academy of Kerala

As we confront one of the world's worst pandemic, our society is steadily moving towards adapting the new normal set by Covid-19. While day-to-day life of many are far from normal, one of the key enablers for humans to survive and also to sustain their livelihood has been the accelerated adoption of Digital Technologies in the last six months.

1. Pre Covid-19 Scenario for Information Technology

The Information technology has been seeing a huge disruption due to the Industry 4.0 developments, where there was a strong convergence between the Cyber, Physical and Biological world.

The major technology areas that were getting very high traction were

- (a) Machine Learning and Artificial Intelligence
- (b) Robotics, 3D Printing
- (c) Internet of Things
- (d) Cyber security
- (e) Wearable technologies
- (f) Augmented Reality
- (g) Robotic Process Automation (RPA)

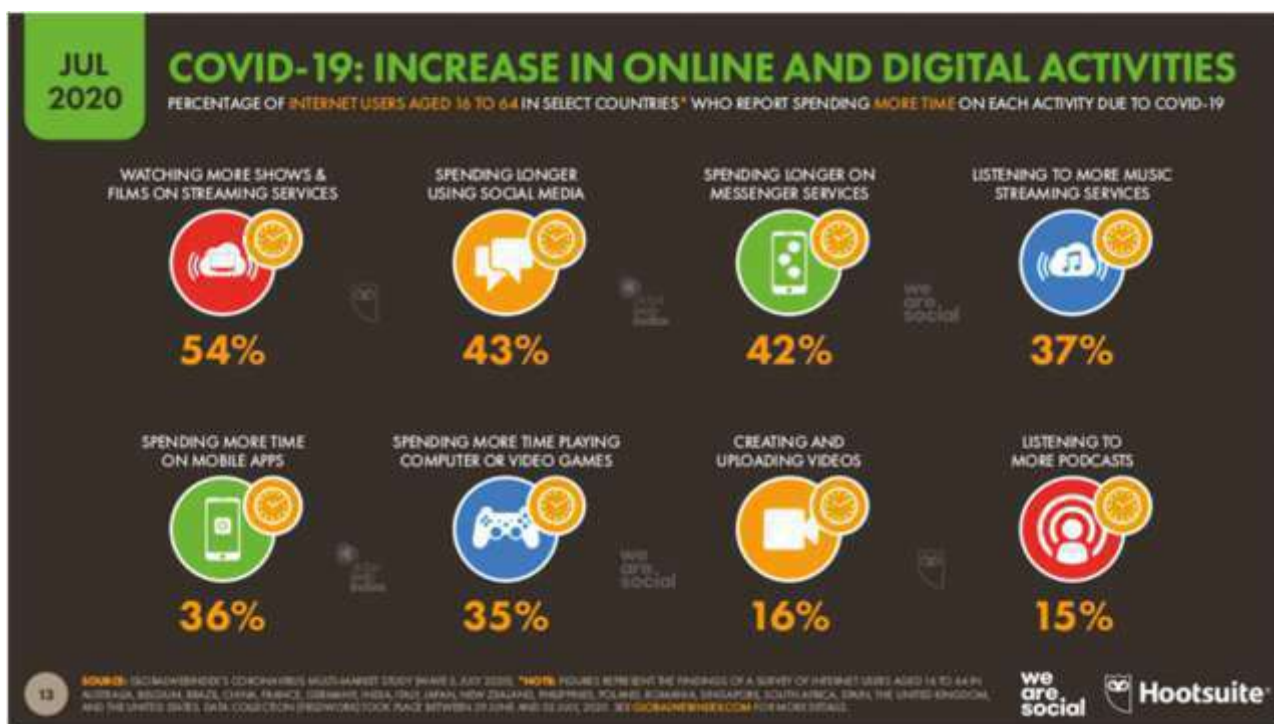
Continuous innovations have led to reduced cost of IT Components and systems. Thanks to providers such as Reliance Jio, the Data and Telecom connectivity world saw competitive pricing making India as the world's most affordable per GB pricing (approx. \$0.09). The data Centre services moved away from on-premise computing to Cloud based

computing services, reducing the average cost by approximately 79 %. Thanks to Chinese manufacturing capabilities, the sensor costs showed a decline from \$1.30 in 2004 to \$0.38 in 2019, leading to a lot of product ideas getting fulfilled in the IOT space. Entrepreneurship ecosystem had a substantial uptick between 2015 and 2019 in the country, and India emerged as the third best startup ecosystem in the world. In terms of smartphone user base, India emerged as No 2 country (after China), which provided a huge potential for S/W apps in the Indian Market. Putting all the information together, world in General and India in particular was poised to be a major player in the information Technology Business. NASSCOM predicated approximately 7.1 % Y-o-Y growth for the IT Sector for the FY 2021, before Covid-19.

2. Trends during Covid-19 Scenario

While most of the Industry sectors except Health were scrambling to stay afloat, Information and Communication Technologies have emerged as an important backbone in solving some of the current crisis situations. Due to the prolonged lockdowns across the globe, use of digital services such as e-commerce, teaching, entertainment or e-governance have increased the digital presence of a large number otherwise novice users. It has been estimated that the data usage during Covid-19 has increased by around 47 % and the internet usage increased by around 40 %.

The Digital 2020, July Global Statshot report by Hootsuite, shows an increase of various common digital services



The increased digital presence of public has opened up avenues for Digital Marketing substantially. Increased digital access for people, however, have posed new challenges in the cyber world for governments, institutions and individuals. Cyber Security investments have gone up and so are the demands in the area.

Use of Data Analytics, especially for Governments across the world has been a widely seen use case. Building prediction models and assessing the potential impact of the disease amongst various strata of people, and thus providing enormous decision-making powers to the operations team on the ground has been very useful for Health departments of various governments. Usage of position tracking applications developed by government of India, for example is of immense use in containing the spread of disease through enabling technologies. These applications are also used as information sharing and awareness development amongst citizens. Even though Robotics has always been a potential enabler for humans in various services, it has been a niche before Covid-19. There were companies who had predicted mediocre growth projects until 2022. However, that trend got an interestingly high

acceleration during the Covid-19 time and hospitals started deploying both supervised and unsupervised Robots for various services, even in remote towns in the country. Use of audio-visual technologies for remote medical consultation and contactless monitoring of patients have been another trend seen in the Healthcare sector. 3D Enabled Respirators and masks helped a large country like India to fight the pandemic effectively, paving ways for the acceleration of Additive printing technologies and pushing them to mainstream.

Customer Experience are being taken to next level by technologists around the world by providing near real life Augmented and Extended Reality based interaction to customers. These technologies are used not only in gaming field, but in other business domains by providing better views for many products and services through the internet. With digital services picking enormous traction, organizations are scrambling through in moving more services online and building more real-life customer experience - web and mobile development are at all-time high for such movements especially small and medium business entities.

Edu-tech has been another interesting domain that used audio / video and collaboration technologies for the institutions to connect their teachers to Students and parents. As per one of the reports of Bloomberg, Google classroom application usage doubled during the peak Covid-19 Lockdown times. Most of the contents were moved digital to help students learn subjects remotely. In an interesting development, the state of Kerala, in India used a combination of online and DTH based transmission facilities to reach to around 1.3 M students during the new academic season that started in June 2020.

With the world economy struggling to pick up and worldwide recession hitting businesses hard, Process automation is seen as an immediate priority for Organizations, as their survival strategy. Companies are cutting operations cost on average by 35 % (likely to go higher if the situation continues) and Robotic process automation is a significant step towards achieving the operations goals.

1. Post Covid-19 New Normal and Skills need

It is predicted that the world will never be the same. At least for a certain period of time, Social distancing in public as well as workplace, constant sanitizing and wearing facemasks in public will remain as the common property for individuals. Remote working, online education, remote health care capabilities will continue to become normal as part of the operations. Industry sectors such as travel, Hospitality, Oil & Gas, Luxury goods, Automobile, Real estate would take more time to pick up, whereas Health care, Technology sectors would bounce back to pre-Covid scenario soon after the restrictions are eased out.

The job market is in turmoil and many countries in the world has seen the highest unemployment rate of the past few decades. In India, Covid has pushed the unemployment rate to 24 %, which was hovering around 6 – 7 % in January 2020. MSME sectors and unorganized sectors suffered a lot more contraction during the Covid-19 period. According to the Q1 performance analysis of 747 small businesses by Care Ratings, their net sales saw negative growth of minus 66.7 per cent from Q1 FY20. Approximately

43 % of MEMSE businesses are facing challenges of shutting down, according to the survey conducted by the All India Manufacturers' Organisation (AIMO)

The main streaming of many IR 4.0 technologies began at least three years ago and with the pandemic scenario becoming an accelerator for digitization of services across many industry domains, it is believed that the job opportunities for youngsters in the information and Communication Technology sector maybe less in the short-term, because of companies being cautious of their future business , it will continue to be the biggest job provider for youths in the country. Full Stack Development, Robotics Process Automation, Data Science and Analytics, Machine Learning / Artificial Intelligence, Cyber Security, Digital Marketing, Extended Reality (XR) are predicted to pick up steam and will remain as top performing IT Skills for the next two to three years. With the new education policy kicking in, and the new direction taken by All India Council for Technical Education (AICTE), many of these IR 4.0 technologies will find a place in syllabus, for higher education. Apart from that, Government schemes may drive more affordable skills development opportunities for the youths leading to a higher quality supply of the ICT skills for future.

Emergence of Freelancing will be a notable trend for jobs in India. Even though such self-employment activities are common in the western world, the trend has been a norm only for the blue-collar jobs in India. With remote working possibilities are time tested, thanks to Covid-19 scenarios, it is anticipated that Freelancing jobs will become a new job type, in the market and the permanent jobs may shrink as the trend gets accepted both from the demand and supply side. Lot of challenges including legal, regulatory and security of information needs to be ironed out on this new job type, it is expected that Freelancing is likely to grow in India. HR technology solutions firm PeopleStrong estimates that 25-30% of workforce in sectors such as internet business companies, IT, ITES, Startups, hospitality, quick-service restaurants (QSR), retail and logistics will get converted to freelance roles over a period of time.

4. References

- ♦ <https://www.cable.co.uk/mobiles/worldwide-data-pricing/>
- ♦ <https://www.statista.com/statistics/748053/worldwide-top-countries-smartphone-users/>
- ♦ <https://www.sherweb.com/blog/cloud-server/total-cost-of-ownership-of-servers-iaas-vs-on-premise/>
- ♦ <https://www.supplychaindive.com/news/declining-price-iot-sensors-manufacturing/564980/>
- ♦ <https://www.marshall.usc.edu/faculty-research/centers-excellence/center-global-innovation/startup-index-nations-regions>
- ♦ <https://in.pcmag.com/the-why-axis/136614/data-usage-has-increased-47-percent-during-covid-19-quarantine>
- ♦ <https://www.indiatimes.com/technology/news/india-internet-usage-increased-by-40-percent-in-covid-19-lockdown-says-report-512955.html>
- ♦ <https://www.financialexpress.com/market/stock-market/care-ratings-ltd-stock-price/>

Global Trends towards Online Trading in Coexistence with Emerging Products and Services

Dr. Swapna K Cherian, Asst.Professor in Commerce, MSM College, Kayamkulam, Kerala

Dr. John T Abraham, Asst.Professor in Computer Science, Bharata Mata College, Kochi

Abstract

Online Trading is the process of buying and selling of goods and services by making use of internet for various purposes related to trade. It also includes transfer of money and data to execute business transactions globally. The retail market in the world has undergone a major revolution and has witnessed remarkable growth in the last few decades. Using different types of e-commerce techniques in various countries help the growth of economy and attain a high level of customer satisfaction in spite of many regulations influencing electronic commerce applications. This paper demonstrates how the trading system influence small, medium and large scale traders including exports and imports with e-commerce applications. It also highlights an overview of e-commerce or internet commerce in the changing business scenario.

Keywords

Online Trading, Economy, E-Commerce, Global Commerce

Introduction

Global Commerce Initiative is a group of manufacturers and retailers that works to simplify and enhance global commerce and improve consumer value. Electronic commerce plays a considerable role in the reducing transaction costs, growing trade opportunities and improving service delivery. In spite of tremendous growth in physical trade transactions, the internet has transformed business operations among traders much easier. The advantages created through the usage of internet in the business world is uncomparable.

With increasing requirements of e-Commerce industry, every businessman is looking to have an online store like that of in Amazon, flipkart, snapdeal, paytm etc. where companies can sell their multi range of products and services. Customer can get enormous benefits by opting online trading as it delivers a wide range of benefits to retailers, merchants and all type of traders irrespective of their product differentiation. E-commerce emerged as a result of the use of the internet and similar technologies. In the fast developing modern world, e-commerce is the most important tool to any strong economy for progress. It is the way of success to businessmen and service sector because it can increase the revenue and provide services more than actual functioning. Normal trade practices can be costly because as it includes lot of other expenses. Using different types of e-commerce techniques in the market help the growth of economy and will achieve a high level of customers' satisfaction even if there are many regulations that affect electronic commerce.

Types of E-Commerce

There are different types of E-commerce that exist with trade and trade related transactions. One of them is B2B. It is called as business to business which means both sides in e-commerce are business or traders. Volume of goods and products are used from one business to another. For example for B2B e-commerce is a small scale industry buying goods or services from a large scale through online. Another type of e-commerce is B2C where by it means that the transactions are between business and consumer. It is the most general form of business that exists and may not have that big

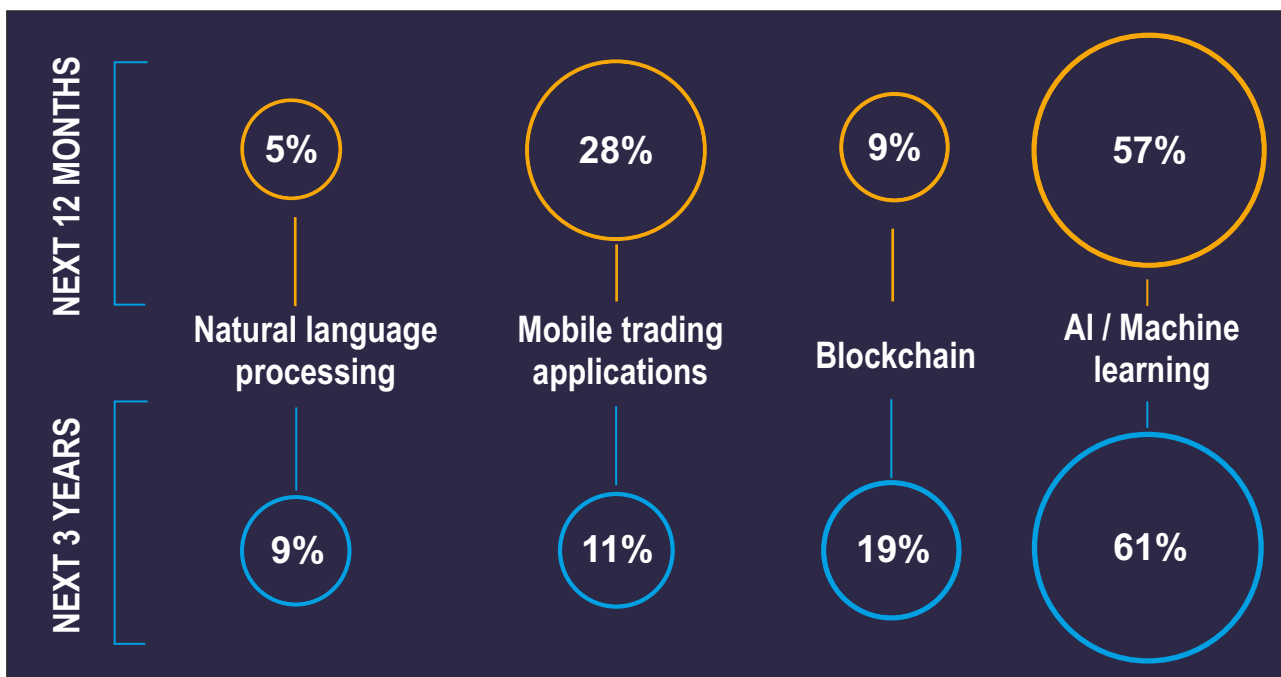
volume for each operation. Normally it has large number of transactions. An example for B2C is a consumer buying goods from Amazon, Flipkart through online marketing source. Yet another type of e-commerce is C2B e-commerce where transaction is between consumer and business. In this type of e-business, a consumer provides such services or requirements to the business people. When a consumer make use of his qualifications, skill and ability for a job, he does this type of e-commerce.

Next type of e-commerce is C2C e-commerce, in which transaction is between consumer and another consumer. Here any consumer or user can sell his products or services to another consumer who is in need of it. The most common type of this type is e-Bay e-commerce. In fact, the www.E-bay.com is a business e-commerce so that we can also call this type of e-commerce as C2B2C

(consumer to business to consumer). In spite of the above common methods of e-commerce there are some other types of e-commerce like B2E which means business to employee, G2G which means government to government, G2E which means government to employee, G2B which means government to business, B2G which means business to government, G2C which means government to citizen, C2G which means citizen to government.

Technologies that helps in the future Global Trading

J. P. Morgan conducted an online survey of e-trading trends with the addition of macro market trends and transformational technologies for 2019. Institutional traders shared their insights and predictions, with the majority being FX traders, the rest being rates and commodities traders.

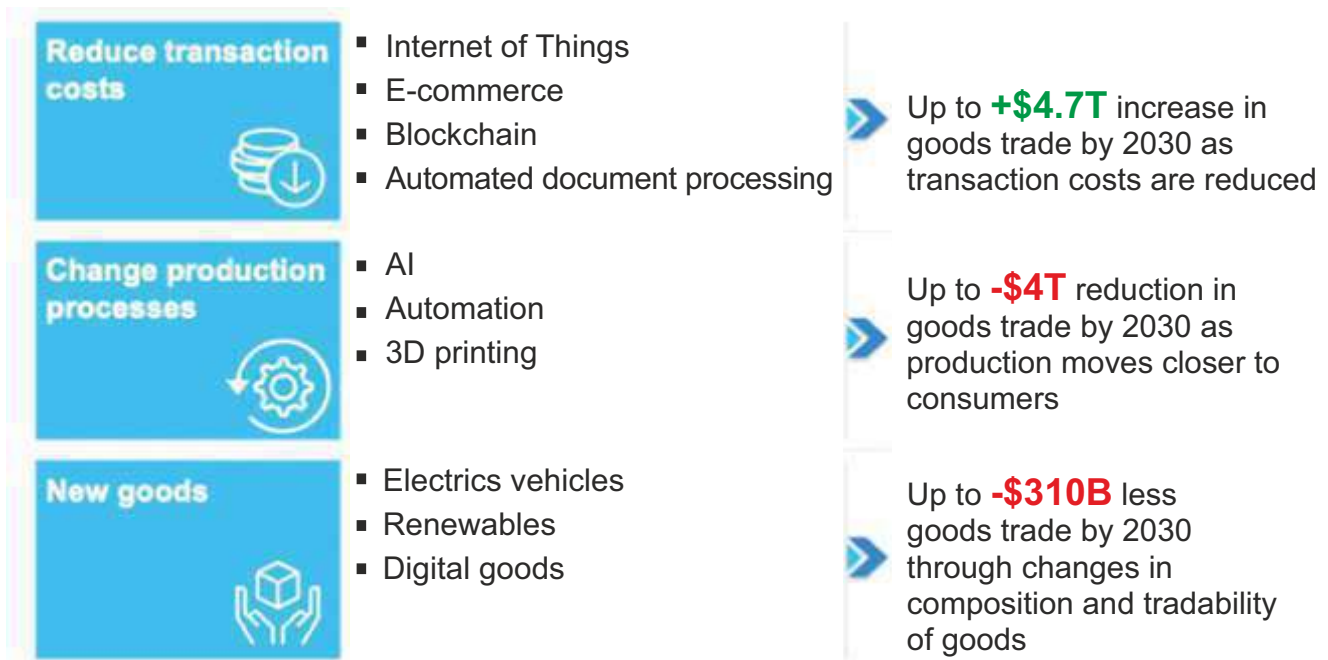


Source : <https://www.jpmorgan.com/global/markets/etrading-trends-2019>

We cannot limit this technological types and their evolvments in trading because of the ongoing progress towards industry 4.0 where digital platforms change their style, look and feel in a regular fashion for attracting customers and retain them permanently. Technological globalization is speeded in large part by technological diffusion, the

spread of technology across borders while the diffusion of information technologies has the potential to resolve many global social problems, it is often the population most in need that is most affected by the digital divide. As always, the history of trade reflects the ongoing march of technological innovation.

Technology is reshaping trade flows in value chains in three ways



Source: McKinsey Global Institute (2019)

Today, Services trade is growing 60% faster than goods trade overall. Some types of services, such as IT services, telecom, business services, and IP royalties, are growing 2-3 times as fast as goods trade. 5G wireless networks, virtual reality, and augmented reality may all give a boost to services in the future by their potential capabilities of enhancements in implications of business processes.

Booming values of E-Commerce

Time management is the main advantage of e-commerce as it saves the valuable time of common man. Time is money and it should be managed perfectly. E-commerce can save the time, minimize the crowd, and make the distance of purchase to nil. Internet and high speed technologies makes buying and selling easier and perform multi tasking at same time in a single place. Through the one computer or mobile screen, the buyer or consumer can browse and search for large number of items as the seller has the opportunity to show the variety of goods and products in the same screen.

Through the webpage, the shape, style, size and

color of products can be shown which can attract the consumers as it gives them an advantage to look and choose their favourite products. Online shopping also provides an additional help to those people who cannot go to the physical market to buy due to the changes in weather and the distance of shop. Some other cases, people are traveling around the world for religious, tourists, professional or business purposes. They may not be able to provide their family or friends a gift and their concern for a special events like birthday etc. as they are far away. By online trading it is easy for them to make the orders any time and any place and enable the person to get the product in their absence. Another benefit is that e-commerce can help to reduce the usage of papers. Papers are expensive and hence its wastage can be reduced with electronic bills.

E-Commerce allows the speedy delivery of goods. It is encouraging competition and thus customer can get lower prices by seeing different products in the online markets. E-Commerce is one of the ideal ways of shopping in the modern world as people enjoy their online trading because of its easiness and convenience. Customers can buy products or

services from their home at any time irrespective of day or night. Purchase can be quick, convenient and user-friendly with the ability to transfer funds online. Because of its convenience, consumers can save their precious time and money by searching their products easily and making purchasing online. Customers get details of the product they intend to purchase from the online product catalog. These information help them to purchase the exact product they wish to buy at any time of the day. The peculiarity of the product and the offers are also made known to the consumers.

In addition to the above benefits the following advantages also exists

- ♦ Allow Happy Customers to Sell Your Products:
- ♦ Boost Brand Awareness:
- ♦ Ability to use Multi-site:
- ♦ Attract New Customers with Search Engine Visibility:
- ♦ Comprise Warranty Information:
- ♦ Decreasing cost of inventory Management:
- ♦ Keep Eye on Consumers' Buying Habit:
- ♦ Competence:
- ♦ Selling Products Across the World:
- ♦ Stay open 24*7/365:
- ♦ Economy:
- ♦ Decrease Costs:
- ♦ Offer Huge Information:
- ♦ Analytics:
- ♦ Expand Market for Niche Products:
- ♦ Scalability:

Conclusion

Besides the enormous advantages provided by online trading it has certain limitations. One of the major draw back is that e-commerce does not provide complete security for the users if they are not cautious regarding their dealings in financial matters. Global commerce is one of the best innovations which have conquered the manufacturing, industrial, operational and service sectors. Moving together with electronic commerce can help us to become a global village by reducing the distance between nations. Encouraging

competition, Easy reach of products, faster services and constant contact with customers' needs are the additional benefits of global commerce. Moving along with the changing trends in the global commerce by making use of online trading is the ultimate requirement of all sectors.

References

1. Keeling, K., L. A. Macaulay, and P. McGoldrick. "Ditv And E-Commerce Among Disadvantaged Community Groups." *Behaviour & Information Technology* 26.6 (2007): 545-560. Academic Search Complete. Web. 27 Jan. 2014.
2. Yanhui, Li, and Zhu Siming. "Competitive Dynamics Of E-Commerce Web Sites." *Applied Mathematical Modelling* 31.5 (2007): 912-919. Academic Search Complete. Web. 30 Jan. 2014.
4. Chan, Felix W. H. "In Search Of a Global Theory of Maritime Electronic Commerce: China's Position On The Rotterdam Rules." *Journal of Maritime Law & Commerce* 40.2 (2009): 185-202. Academic Search Complete. Web. 28 Jan. 2014.
5. Molla, Alemayehu, and Richard Heeks. "Exploring E-Commerce Benefits for Businesses In A Developing Country." *Information Society* 23.2 (2007): 95-108. Academic Search Complete. Web. 29 Jan. 2014.
6. Yanhui, Li, and Zhu Siming. "Competitive Dynamics Of E-Commerce Web Sites." *Applied Mathematical Modelling* (2007): 912-919. Academic Search Complete. Web. 30 Jan. 2014.
7. Goldfarb, Avi. "The Internet Killed Distance. Mobile Phones Brought It Back." *Technology Review* 117.1 (2014): 62-63. Academic Search Complete. Web. 27 Jan. 2014.
8. <https://www.ukessays.com/essays/information-technology/e-commerce-and-online-trading-information-technology-essay.php>
9. <https://medium.com/@briannawillsss/18-major-benefits-of-e-commerce-business-for-retailers-customers-in-2015-63c5fc87f679>
10. <https://voxeu.org/article/next-generation-technologies-and-future-trade>

Curriculum and Mentoring Changes required in Primary and Secondary Education in the era of Industrial Revolution 4.0

Dr V T Gopakumar

In the fourth industry revolution (I4.0), we are experiencing a range of new technologies that combine the physical, digital and biological worlds. These new technologies will impact all areas of our life which in turn has an impact on economies and industries. These revolutionary technologies connect billions of people to cyber physical systems (CPS) and other interconnected networks. This will definitely improve the productivity, flexibility, as well as enhanced product and service qualities of businesses and organizations. It is reported that the I4.0 would be more witnessed in manufacturing industry, as it offers the opportunity for manufacturers to optimize all their operations very quickly and efficiently. In the manufacturing industry I4.0 is working on the concept of smart factory; where the equipment and machines are integrated with workers through CPS. The main technologies that enable I4.0 include smart sensors, automation circuits and devices, advanced and sophisticated robots, Internet of Things, cloud computing, location detection Technologies, human-machine interface, augment reality, 3D printing, artificial intelligence, big data analytics etc...

In the era of I4.0, the students graduating from various technical and non technical institutions must have the necessary skill set and work force pre-requisite with the I4.0 technologies. A few of them are; skills in automation, digitization, and information technology. They must be ready to do embark on a path of continuous learning and adaptation to meet the new industry requirements rather than rely on the knowledge from the traditional education system alone. Here in I4.0 we are witnessing the overlap of man and machine and

it is essential to have exposure to humanities, sociology and soft skills in addition to the traditional technical curriculum built around science and innovation imparted to us through our primary and secondary education system. At the same it is much more important that the teachers/faculties/mentors must have the competence to train these young minds intelligently as per the skill and knowledge requirements arising from time to time.

Here are some of the views to up-skill our students, especially since their current schooling system does to prepare them to succeed in the era of I4.0. They would need build expertise or at least get an exposure to the intellectually intensive jobs. Yes, we need to act as quickly as possible to do what is needed to bring about a revolution in the current education system as well. Better to label it as Education 4.0 (E4.0) for our young minds. It has been reported that technologically advanced countries already have taken necessary and adequate steps to avoid unplugging the youngsters because of the lack of adequate skill set requirements for I4.0.

I would like to put forward a few suggestions to implement in our current primary and secondary education system, so that the undergraduate and graduate students who graduate from the technical and non technical institutions can easily cope up with the skill sets requirements for the coming years of I4.0. Inculcating these in the early years, at the primary and secondary school levels could train them and provide them the necessary skills to be eligible for job requirements and demands came across.

Some steps that I would suggest to help enable this are as follows. Give computer education as part of their primary education. Offer them introduction to electronic circuits and computer operations of their age, make them familiar with basic computer skills as word typing, Microsoft office packages as Excel to do simple mathematical operations and puzzles, hobby electronics circuits and watch educational YouTube videos under qualified faculty supervision. In higher classes they can be introduced to free and open source software later to do simple experiments using low cost robotic kits like arduino board and raspberry pi. Certainly this would enable our children to develop a love of learning their demanding lessons through hands-on experiences that highlight their collaboration and creativity. Students must be mentored to go beyond their traditional curriculum and go through different learning methods. This will also introduce them to necessary corporate soft skills as organizational skills, time management skills and collaborative skills, which in turn would help them not only with their academics but also employment in future. The Government and policy makers under the government should make sure that all schools are furnished with computer lab facility centers with enough number of PCs and required software packages for the students. Also it would be ideal for the government to provide at a subsidized rate, low cost educational PCs for each student. Additionally, at least once a month, the students could be exposed to invited lecture and experimental lab works from industry professionals. In addition to that, it is essential to change the traditional Q&A or subjective writing examination as it may not suffice the needs of future talent requirements of the industry. The assessments, as part of Education 4.0, will not be solely based on the current examination patterns. It will be done by analyzing their learning journey through practical and experiential learning-based projects or field works. Also taking in to account that, the mentors must play a vital role to create an ideal mindset for the young graduates, it is necessary to have industry mentoring programs introduced in the curriculum for re-skilling and up-skilling.

In the era of I4.0 there are more human machine interface, replacing the human by smart machines and it is inevitable. While doing so, the I4.0 technological developers must have more imagination, critical thinking, creativity, writing ability, presentation skills, negotiation, curiosity, spiritual and moral values and ethics. Emotional intelligence, leadership and social influence as well as service orientation also see an out sized increase in demand relative to their current prominence. This means humanities training are to become inevitable in the primary and secondary education system. It has to be noted that future employees would be global team players; they must have a global ethos. Similarly proficiency on more than one or two international language would certainly be an added advantage.

While considering I4.0 readiness in graduation studies, the current existing skill gap should be nullified for taking new jobs in I4.0. The World Economic Forum's - The Future of Jobs report has predicted that around 75 million jobs worldwide will be displaced by automation by 2022. But in the same period, 133 million new jobs will be added, but the very unfortunate fact is that the workers who lose their present jobs will not have the necessary qualifications or the adequate skill sets to fill the new job positions created. The present education system follows four years of study to qualify for a degree needed. In the graduate studies too there must be necessary changes made to address the new industry demands and requirements. This could be facilitated by online courses (education at anytime and anywhere), certificate programs offered by reputed organizations and universities, internships at reputed firms and of course become the part of startups ventures and entrepreneurship. If the educational institutions have valuable intra - and inter-industry partnership and collaboration, the institutions can reorient their curriculum as per the demands of the industry on one hand, thereby benefiting the industry on the other hand as they can easily pick students for internships and offer job thereafter. Hence learning institutions provide employable students rather than just educated students who can be moulded by the industry.

In conclusion, our new generation has to prepare themselves to address the new jobs arising. Among the range of roles that are expected to experience increasing demand in the period up to 2025 are Data Analysts and Scientists, Software and Applications Developers, and e-commerce and Social Media Specialists. Also it is expected that there would be a growth in roles that leverage distinctively 'human' skills such as Customer Service Workers, Sales and Marketing Professionals, Training and Development, Organizational Development Specialists as well as Innovation Managers. There is extensive evidence of accelerated demand for a variety of wholly new roles as Artificial Intelligence (AI) and Machine Learning (ML) Specialists, Big Data Specialists, Process Automation Experts, Information Security Analysts, User Experience and Human-Machine Interaction Designers, Robotics Engineers and Blockchain Specialists etc....It is high time for national policy makers to implement policies to build an ecosystem of well-skilled workforce and of national reskilling and upskilling equipped to support our workers of all ages to keep abreast of technological change and develop skills in need. The governments need to address the impact of new technologies on labor markets through upgraded education policies aimed at rapidly raising education and skills levels of individuals, especially in STEM (science, technology, engineering and mathematics) and non-cognitive soft skills, enabling people to strengthen their uniquely human capabilities. The government's intervention must include changes in school curricula, teacher training and a reinvention of vocational training for the age of the Fourth Industrial Revolution. Equally important that, changes and improvements in education and skills preparation or implementation

must be balanced equally with the demand side. All these demands a revolution in education to enable people to harness the opportunities created by the advent of the above emerging technologies mentioned. Finally the school education must look and foresee these job requirements of future and change curriculum accordingly and train our children for their better future and for our country as well. Adding to that, the premier institutions, institutions of national importance corporate industries must come forward and give their recommendations to policy makers to meet the skills and knowledge requirements for our children in the new industry revolution.

References

1. <https://www.weforum.org/agenda/2013/11/can-start-ups-help-higher-education-to-shape-up>
2. <https://www.weforum.org/agenda/2019/01/how-students-can-graduate-qualified-for-fourth-industrial-revolution/>
3. http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf
4. <https://www.futurereadyedu.com/what-is-education-4-0-how-you-can-adapt-this-in-the-learning-environment/>
5. <https://www.weforum.org/agenda/2018/09/future-of-jobs-2018-things-to-know/>
6. <https://www.weforum.org/agenda/2019/04/skills-jobs-investing-in-people-inclusive-growth/>
7. <https://www.vinnter.se/the-value-of-robotic-automation-lies-in-reskilling-the-workforce/>

Banknote Recognition for Visually Impaired using Machine Learning

Dr. Nijil Raj N

Prof. and Head, Dept. of Computer Science and Engineering,
Younus College of Engineering and Technology, Kollam, Kerala, India
Email: csithod.nijil@ycet.ac.in

Anandu S Ram

anandu.s.ram@gmail.com

Aneeta Bino Joseph

aneetabino@gmail.com

Shabna

shabnashameer27@gmail.com

B.Tech Students, Department of Computer Science and Engineering,
Younus College of Engineering and Technology, Kollam, Kerala, India

***Abstract-**One of the most significant issues faced by visually impaired people is the recognition of paper currency. The current system uses Malaysian Ringgit bank notes, the feature set used is RGB values and the algorithms used are KNN, SVM, Naive bayes and Decision Tree. In our proposed system there are three phases. In phase I four features are extracted and in phase II RGB values are concatenated with phase I and finally in phase III, phase I and phase II are concatenated. Different ML methods are used such as KNN, Decision tree, SVM and Naive bayes. In our system svm provides a better accuracy of 100%.*

***Index Terms -** Banknote Recognition, Machine Learning*

I. INTRODUCTION

Identification of banknotes is a relatively easy job for the human with normal eyes as the brain is capable of absorbing different information and recognizing them with less effort. But for visually impaired people, banknote identification has become one of the troublesome things to be faced as a result of the quality of the banknote itself. The visually impaired are divided into three subgroups: colour-blind, partially sighted and blind people. Their basic necessity is to identify the denomination rather than security measures, as they assist them in determining the type of banknote. Up to this point, the almost no distinction in size of banknote and an effectively blur away material blemish on the outside of banknote consistently offer trouble to visually impaired individuals to identify and characterize bank note appropriately.

The World Health Organization reports states, 285 million people worldwide are calculated to be visually impaired which includes 39 million people

totally blind and 246 million people have low vision. Visual impairment is defined as severe minimizing in vision that cannot be corrected with the contact lenses or even standard glasses and reduces a person's ability to perform certain or all tasks. Visual impairment may cause people to have difficulties and limit their activities of daily living such as doing housework, cooking, shopping, watching television, reading, walking, socializing with other and so on. Therefore assistive technology is needed to help them to overcome these challenges. Our system recognizes Indian currencies for visually impaired people. It uses 4 algorithms namely, k-nearest neighbour, decision tree classifier, support vector machine and Bayesian classifier.

In the existing system the feature set selected is RGB values. Our project is divided into three phases such as Phase I, Phase II and Phase III. In phase I our features are considered namely, variance, kurtosis, entropy and skewness. In the phase II RGB values are considered and in the last phase RGB values and the four features are considered.

II. LITERATURE SURVEY

1. N. A. J. Sufri, et.al., propose a vision Based System for Banknote Recognition Using Different Machine Learning and Deep Learning Approach. They used the rgb values as features and used algorithms DT, NB, KNN, SVM and deep learning alexnet. Both kNN and DTC achieved 99.7% accuracy but both SVM and BC perform better by succeeded to achieve 100% accuracy.

2. Kamal, et al., proposed feature extraction and identification of Indian currency notes. The work separates particular and remarkable highlights of Indian currency notes, for example, central numeral, RBI seal, color band and ID mark for the visually impaired and utilizes calculations improved for the recognition of every particular component. The strategy has been assessed over an enormous informational index for acknowledgment of Indian bank notes of different divisions and states of being including new notes, wrinkled notes and non- uniform light. Intensive examination yields a high true positive rate (wanted element distinguished accurately) of 95.11% and a low false positive rate of 0.09765% for emblem recognition, a precision of 97.02% for central numeral location, and 100% correctness for both acknowledgment of recognizable ID mark and color matching in CIE LAB colour space affirming the proficiency and strength of the proposed conspire.

3. Gouri Sanjay Tele, et al., proposed detection of Fake Indian Currency. Security highlights of currency are basic for deciding genuine and fake money. Basic security highlights incorporate watermarks, dormant pictures, security thread, and optically variable ink. In this a methodology for counterfeit currency location extricates the general traits of latent pictures and distinguishing ID mark from the image of money. Extricate properties from images of currency notes can get very intricate as it includes the extraction of some noticeable and undetectable highlights of Indian currency. After demonetization 500 and 2000 are the high esteemed cash notes existing till date so there is a most extreme likelihood that this notes can be duplicated so as to maintain a strategic

distance from this they use programming to identify the fake notes utilizing picture handling procedure.

4. Navya Krishna G, et al., proposed Recognition of fake currency note using CNN. The Automatic Fake Currency Recognition System (AFCRS) is intended to identify the fake paper money to check whether it is fake or original.

The current fake issue because of demonetization impacts the financial framework and furthermore in different fields. Another methodology of Convolution Neural Network towards recognizable proof of fake notes through their images is inspected in this paper which is relatively better than past image processing strategies. It depends on Deep Learning, which has seen huge accomplishment in image classification lately. This procedure can support both people and machine in recognizing fake notes progressively through an image of the equivalent. The proposed framework, AFCRS can likewise be conveyed as an application in the smart phone which can assist the general public with distinguishing between the original and fake notes. The Accuracy in the undertaking can be expanded through the original fake notes.

TABLE I

COMPARITIVE ANALYSIS OF LITERATURE SURVEY

Method	Existing System Accuracy
SVM	97.02%
CNN	-
CNN VGG Net	80%

The literature survey concludes that there are many disadvantages with the existing systems. Our proposed method of banknote recognition extract features from bank notes using Machine learning algorithms such as, KNN, DTC, SVM and BC. The Comparative Analysis of our literature survey is shown in Table I

III. MATERIALS AND METHODS

A. Dataset

The data set used is Indian currency. The data set contains various Indian currencies of Rs20, Rs50, Rs100, Rs200, Rs500. It has various security features of Indian currency.

B. Feature Set

The features selected are:

1. Entropy : Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image. Entropy, h can also be used to describe the distribution variation in a region. Overall Entropy of the image can be calculated as:

$$h = - \sum_{i=0}^{K-1} Pr_i (\log_2 Pr_i) \quad (1)$$

where, Pr is the probability of the i -th grey level, which can be calculated as $Z_i / m \cdot n$, Z_i is the total number of pixels with the i -th grey level and K is total number of grey levels.

2. Skewness : Skewness characterizes the degree of a symmetry of a pixel distribution within the required window around its mean. Skewness may be a pure number that characterizes only the shape of the distribution. The formula for finding Skewness is given in the below equation:

$$S = \frac{1}{m \cdot n} \sum \sum \frac{(p(i,j) - \mu)^3}{\sigma^3} \quad (2)$$

where, $p(i, j)$ is the pixel value at point (i, j) , m and n are the mean and standard deviation respectively.

3. Kurtosis : Kurtosis measures the Peakness or flatness of a distribution relative to a normal distribution. The conventional definition of kurtosis is:

$$K = \frac{1}{m \cdot n} \sum \sum \frac{(p(i,j) - \mu)^4}{\sigma^4} \quad (3)$$

where, $p(i, j)$ is the pixel value at point (i, j) , m and n are the Mean and Standard Deviation respectively.

4. Variance : Variance is the square root of standard deviation. The formula for finding Variance is:

$$Var = SD^2 \quad (4)$$

where SD is the Standard Deviation.

5. RGB Values : Each of databases were used to extract colour values - red, green and blue intensity average values called as RB, RG and GB using equations(5)to(7):

$$RB = r - b \quad (5)$$

$$Rt = r - g \quad (6)$$

$$tB = g - b \quad (7)$$

whereby r is the average intensity value for red channel, b is the average intensity values for blue channel, and g is the average intensity values for green channel of the pixels inside the cropped region.

C. Algorithms Used

1) K Nearest Neighbour: In pattern recognition, the k -NN may be a method used for both classification and regression [3]. In each cases, the input consists of the k closest training examples within the feature space. The output depends on whether k NN is employed for classification or regression:

In k NN classification, the output may be a class membership.

In k -NN regression, the output is that the property value for the object.

2) Support Vector Machine: They are supervised learning models with associated learning algorithms in ML that analyze data for classification and regression analysis [4]. Given a group of coaching examples, each marked as belonging to at least one or the opposite of two categories, an SVM training algorithm build a model which assigns new examples to one category or the opposite, making it a nonprobabilistic binary linear classifier[5].

3) Naive Bayes classifier: They are a family of probabilistic classifiers on applying Naive Bayes' theorem with strong assumptions within the features. They are among the sole Bayesian network models. Naive Bayes has been studied extensively since the 1960s. it had been introduced in the early 1960s, and remains a popular method for text categorization, the matter of judging documents as belonging to one category or the other with word frequencies because the features. It helps in automatic diagnosis.

4) Decision tree: A tree uses a model like tree of decisions and their possible consequences, including event outcomes, utility. It is a choice support tool. Decision trees are commonly utilized in operations research, especially in decision analysis. A choice tree could also be a flowchart-like structure during which each internal node represents a “test” for each branch represents the results of the test, and each leaf node represents a category label. The path from root to leaf represents classification rules. In decision analysis, a choice tree and thus the closely related influence diagram are used as a clear and analytical decision support tool, where the expected values of them are calculated[6].

D. Standard Measures

1. Precision: Precision is defined as follows:

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

2. Recall: Recall is defined as follows:

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

3. Accuracy: Accuracy can also be calculated in terms of positives and negatives as follows:

$$\text{Accuracy} = \text{TP} + \text{TN} / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

4. f1 Score: It can be calculated as follows:

$$\text{F1score} = 2 * (\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision})$$

IV. METHODOLOGY

Our system is divided into three phases - Phase I, Phase II and Phase III. Phase I extracts features such as variance, kurtosis, entropy and skewness. Phase II extracts RGB values and in the phase III all the five features including variance, kurtosis, entropy, skewness and RGB values are considered. Our data set has rs20, rs50, rs100, rs200 and rs500 notes (5 classes of notes). Then random snaps are taken out of these classes of notes. 15 images of 11 positions are considered.

These positions are at random angles so that when the user inputs the image, the denominations are correctly predicted. The 11 positions are discussed below:

- P1: Image at straight focus (horizontal or vertical)
- P2: Image at backside focus (horizontal or vertical)
- P3: Image at left flip (front side)

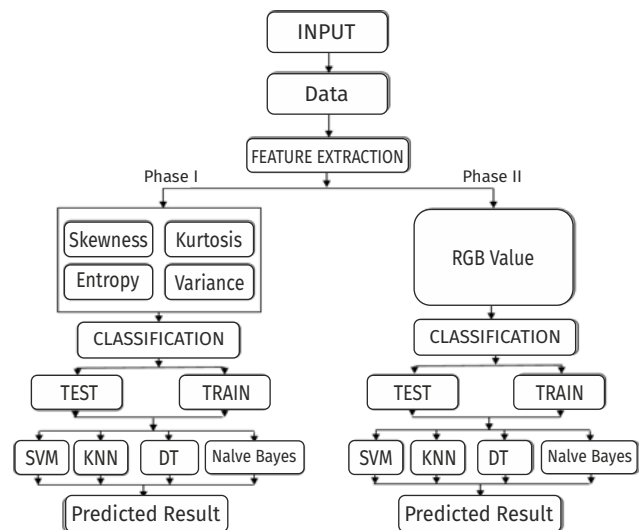


Fig. 1. Phase I & Phase II

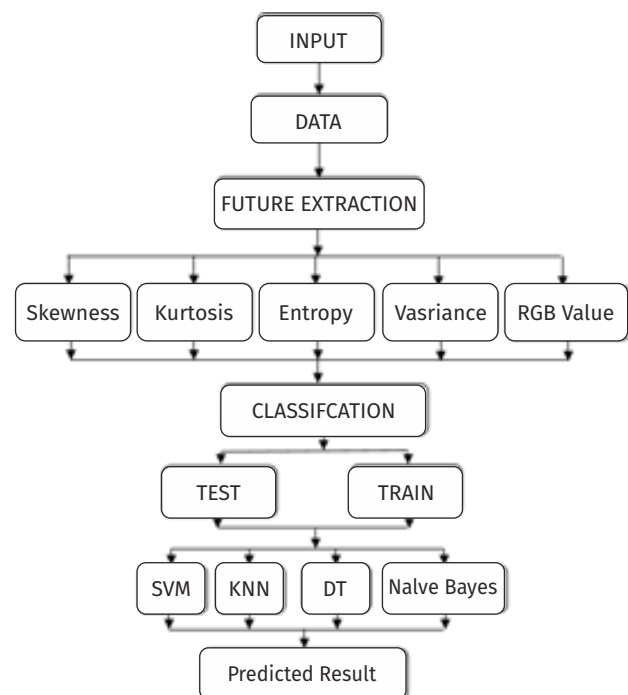


Fig. 2. Phase III

CLASSIFICATION

P4: Image at right flip (front side)

P5: Image at left flip (back side)

P6: Image at right flip (backside)

P7: Image folded to half from upper side at straight focus

- P8: Image folded to half from lower side at straight focus
- P9: Image folded to half from upper side at backside focus
- P10: Image folded to half from lower side at backside focus
- P11: image folded to 1/4th from lower at straight focus.

So from the above there is a total of 165 images for each of the currencies.

A. Classification

Classification represents the problem of identifying to which a set of categories of a new observation belongs, on the basis of training set of data which contains observations whose category membership is known.

B. Test & Train Set

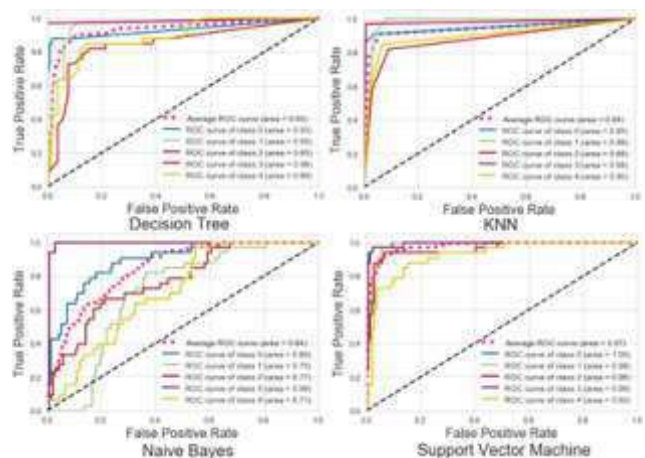
A training dataset is used for learning, that is to fit the parameters for eg, a classifier. A test dataset is a dataset that is independent of the training dataset, but follows the same probability distribution as the training data set.



Fig. 3. Shows 11 positions

C. Algorithm

Finally, Machine Learning Classification model such as k-Nearest Neighbor (kNN), Decision Tree Classifier (DTC), Support Vector Machine (SVM) and Bayesian Classifier (BC) were used to recognize and classify the different values Indian Currency based on the above discussed features.



V. RESULTS AND DISCUSSION

For the first stage the first four features ,i.e, Entropy, Skewness, Kurtosis, Variance were considered and the result is given in table II

TABLE II

PHASE I

Algorithm	F1 - Score	Precision	Recall	Accuracy	
				Training	Testing
KNN	0.861	0.866	0.861	0.950	0.861
Decision Tree	0.802	0.806	0.800	0.936	0.800
SVM	0.731	0.763	0.745	0.736	0.745
Naive Bayes	0.487	0.642	0.539	0.526	0.539

The ROC curve for Phase I is shown in fig 4.

For the second stage the rgb values were considered and is given in table III

TABLE III

PHASE II

Algorithm	F1 - Score	Precision	Recall	Accuracy	
				Training	Testing
KNN	0.888	0.891	0.891	0.885	0.867
Decision Tree	0.853	0.852	0.855	0.956	0.855
SVM	0.703	0.842	0.770	0.768	0.770
Naive Bayes	0.781	0.777	0.788	0.774	0.778

The ROC curve for Phase II is shown in fig 5.

For the final stage all the features were considered and hence the result in table IV.

The ROC curve for Phase III is shown in fig 6.

Fig. 4. ROC curves of Phase I

Fig. 5. ROC curves of Phase II

The disadvantage with the existing system is that only less features are considered for recognizing the various currencies. The comparison of accuracies of existing and proposed systems are shown in Table V

In the existing system the algorithms KNN, DT, SVM and NB shows higher accuracy compared to our proposed system because they use Malaysian ringgits banknotes which have less security features compared to Indian currency banknotes. In addition to that the existing system only takes into account.

TABLE IV PHASE III

Algorithm	FL - Score	Precision	Recall	Accuracy	
				Training	Testing
KNN	0.982	0.983	0.982	0.995	0.988
Decision Tree	0.951	0.953	0.952	0.988	0.952
SVM	1.000	1.000	1.000	0.997	1.000
Naive Bayes	0.903	0.903	0.903	0.871	0.903

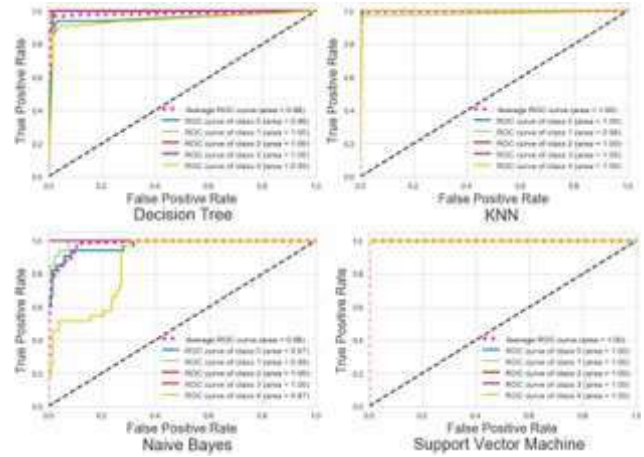


Fig. 6. ROC curves of Phase III TABLE V

ACCURACY COMPARISON OF EXISTING AND PROPOSED SYSTEM

Algorithm	Existing System	Proposed System		
		Phase I	Phase II	Phase III
KNN	100%	86.1%	89.1%	98%
DT	99.7%	80%	85.2%	95%
SVM	99.7%	87.9%	82%	100%
NB	100%	53.9%	79%	90%

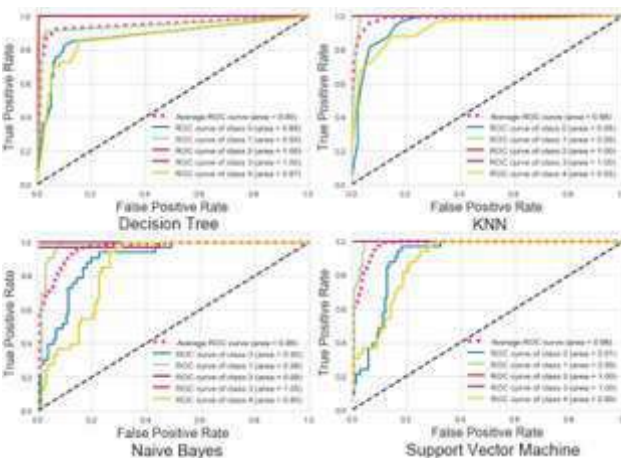
a single feature i.e, RGB values but our proposed system considers 5 features. From the result and discussions we can conclude that our system is better than the existing system.

VI. CONCLUSION

A vision based automated algorithm that can recognize and classify Indian Currencies using machine learning were well developed. svm was found to give best results. From this algorithm, the visually impaired people able to improve their quality of life by reduce the dependency to other especially during outside activities.

VII. ACKNOWLEDGMENT

We take this opportunity to express our sincere gratitude to all those without whom this project would not have been a success. First of all, we owe



our thanks to the Almighty for providing us the strength and courage to complete the project. We express our deep and sincere gratitude to our guide Dr. Nijil Raj N, Head of the Computer Science and Engineering Department, Younus College Of Engineering And Technology for providing valuable advice and timely instructions, without which we could never have been able to complete the work in time.

REFERENCES

1. NAJ Sufri, NA Rahmad, NF Ghazali, N Shahr, and MA As'ari. Vision based system for banknote recognition using different machine learning and deep learning approach. In 2019 IEEE 10th Control and System Graduate Research Colloquium (ICSGRC), pages 5–8. IEEE, 2019.
2. NA Jasmin Sufri, NA Rahmad, MA As'ari, NA Zakaria, MN Jamaludin, LH Ismail, and NH Mahmood. Image based ringgit bank note recognition for visually impaired. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 9(3-9):103–111, 2017.
3. Pascal Soucy and Guy W Mineau. A simple knn algorithm for text categorization. In *Proceedings 2001 IEEE International Conference on Data Mining*, pages 647–648. IEEE, 2001.
4. Baiqing Sun and Jilu Li. The recognition of new and old banknotes based on svm. In 2008 Second International Symposium on Intelligent Information Technology Application, volume 2, pages 95–98. IEEE, 2008.
5. SVM Vishwanathan and M Narasimha Murty. Ssvm: a simple svm algorithm. In *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290)*, volume 3, pages 2393–2398. IEEE, 2002.
6. S Rasoul Safavian and David Landgrebe. A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3):660–674, 1991.
7. Zsófia Solymár, Attila Stübendek, Mihály Radványi, and Kristó fKaracs. Banknote recognition for visually impaired. In 2011 20th European Conference on Circuit Theory and Design (ECCTD), pages 841–844. IEEE, 2011.
8. Kuldeep Verma, Bhupesh Kumar Singh, and Anupam Agarwal. Indian currency recognition based on texture analysis. In 2011 Nirma University International Conference on Engineering, pages 1–5. IEEE, 2011.
9. R Vishnu and Bini Omman. Principal features for indian currency recognition. In 2014 Annual IEEE India Conference (INDICON), pages 1–8. IEEE, 2014.
10. Samina Khalid, Tehmina Khalil, and Shamila Nasreen. A survey of feature selection and feature extraction techniques in machine learning. In 2014 Science and Information Conference, pages 372–378. IEEE, 2014.

Netpreneurship: Motives, Impediments and Satisfaction among Netpreneurs in North Kerala

Dr. G. Kanagavalli

Assistant professor
Department of Commerce
Alagappa University, Karaikudi

Ramseena Azeez

Research Scholar
Department of Commerce
Alagappa University, Karaikudi

***Abstract:** The process of entrepreneurship has been defined as managerial behavior that consistently exploits opportunities to deliver results beyond one's capabilities (Parston, 1998). An entrepreneur is someone with vision who spots a new opportunity and acts on it. The term e-entrepreneurship can be used to mention the process or the work done by individual who is involved in the continuous improvement of their entrepreneurship which are mainly based on social media like Instagram, Face book, Whats App etc.*

The e-entrepreneurship has widened a lot in recent times since the social media involvement of current generation is very vast. Almost all individual out there in current generation is connecting to social media site or mobile application every now and then; this may be the main motive of e-entrepreneurs. The e-entrepreneur can be termed as Netpreneur. It is basically means a small start-up which solely online (net based) with no physical office and their website, blog, e-presence is their office. In this study on netpreneurship, mainly researching on the Netpreneurs and the various data about the Netpreneurship from north Kerala. This study is mainly discussing the various aspects of Netpreneur, such as, various motives of the Netpreneurship, impediments and satisfaction of the netpreneurs in north Kerala, challenges included in the process etc.

KEY WORDS: NETPRENEURSHIP, E- ENTREPRENEURS, NETPRENEURS

1. INTRODUCTION

Netpreneur is an entrepreneur who applies innovation to create new business on internet or someone who publishes his or her business online. The term Netpreneur is a combination of two words "NET" which refers to the internet, which is a global system of interconnected computer networks that use standard internet protocol suite to serve billions of users worldwide and "ENTREPRENEUR" which refers to a person who undertakes innovations, finance and business acumen in an effort to transform innovations into economic goods. Most of the people are now converting their passion into their profession due to the increased development of internet. This trend leads to high economic growth. Most of the women are now entering in this

field due to increased opportunity. This study is mainly discussing the various aspects of Netpreneur, such as, various motives of the Netpreneurship merits and demerits of the netpreneur, challenges included in the process etc.

Netpreneurship

The term "Netpreneurship" is actually formed from two familiar words, one is Net which referred to internet and the second one is Entrepreneurship. It is a business which is based online without any physical office. Netpreneurship stands for electronic commerce and caters to purchase and selling of products and services and information via internet. It is a process of starting a business venture, funding and organizing the required sources and taking both the risk and rewards associated with the venture

online. The new generation of dynamic internet users vastly depends on the online Medias and other internet stuffs for even day to day life, for example ordering food online and ordering stationary items online since the providers are delivering them at the door step.

2. STATEMENT OF THE PROBLEM

The Netpreneurship is an innovative way of thinking the entrepreneurship by utilizing the internet resources. Even though most of the individuals are easily starting the move for the Netpreneurship not everyone is success in the process. Idea of this study is to find the common problems identified by various individuals on the Netpreneurship and suggest solution based on the inputs from people who are well succeeded on this field.

3. OBJECTIVES OF THE STUDY

1. To know the socio-economic and demographic features of netpreneurs.
2. To identify the motivating factors which lead to start Netpreneurship
3. To identify the challenges faced by netpreneur in North Kerala
4. To understand whether there is any association between use of multiple promotion methods and Income from business.
5. To examine the satisfaction level of netpreneur.

4. SCOPE OF THE STUDY

The study on the netpreneur includes the various motives of individuals who are involved in this kind of entrepreneurship. It also comprises of several aspects, challenges and advantages/ disadvantages of these entrepreneurship through internet. The process with which they are utilizing the resources in the internet and the planning on the promotion, futuristic business forecast strategy which they are following may also be included in this study. The various Problems / challenges, and their level of satisfaction etc... are the areas under this study.

5. METHODOLOGY

The study is designed as an empirical one based on the survey method. The samples under the study are taken from the netpreneurs in north region of Kerala state. The study conducted during the month of march 2020. 60 netpreneur are selected as sample size. The study is primarily based on the primary data collected from the respondent. To gather primary data snowball sampling technique is used. The questionnaire was prepared to collect information regarding motives, attitude, Impediments and satisfaction level of netpreneurs in north region of Kerala. North region of Kerala includes the districts of Kasaragod, Kannur, Kozhikode and Wayanad. The secondary data has been collected from various publications, journals, periodical and articles related with the study. Percentage method has been used as a tool for analysis and interpretation of collected data. The collected data statistically analyzed and presented here.

6. LIMITATIONS OF THE STUDY

- The netpreneur who are involved in this study may not be doing the business of same idea, thus the data generalization cannot be done.
- Data collection will be difficult since the person who is handling the Netpreneurship may not easily be available to take the questionnaire session. The collected data will be limited the area – north Kerala and The response from the individuals may not be accurate.

FINDINGS OF THE STUDY

Age	Male	Percentage	Female	Percentage
Below 15	0	0	5	8
15 - 23	2	3	28	47
23 - 31	3	5	18	30
31 - 39	0	0	4	7
Total	5	8	55	92

(Source: primary data)

Mostly youngsters are interested in doing this kind of business. The analysis state that the majority of people who has age group between 15 - 23 are involved in this kind of business endeavors. It is inferred that 92 percentages of respondents are females who involved in Netpreneurship.

Table No: 2
Occupation wise classification (present status)

Present status	Number of respondents	Percentage
Students	35	58
House wife	18	30
Business	4	7
other	3	5

Source: primary data

The netpreneurship is mostly chosen by students. 58 %of respondents are students while 30 % respondents are housewives.

Table No: 3
Most preferred product

Category of respondents	Number of respondents	Percentage
Craft items	20	34
Garments	14	23
Baked food items	24	40
Others	2	3

Source: primary data

Craft items creation and baking food items are mostly preferred netpreneurship. Out of 60 respondents 24 of them are chosen crafting as their way of earning while 20 of the respondents choose baking as their business.

Table No: 4
Motive for Netpreneurship

Category of respondents	Number of respondents	Percentage
Passion	56	93
Profession	4	7
Other	0	0

Source: primary data

Passion is the main motive to the individuals who are interested in this kind of business. 93.33%of respondents choose netpreneurship because of the passion they have towards this business. Only 6.67% has been chosen this as their profession.

Table No: 5
Source of business idea

Category	Number of respondents	Percentage
Social media	32	54
Friends or Family	20	33
Self	8	13

Source: primary data

The table shows that the 54% people got the idea for the netpreneurship from Social media, 33% of them got the idea from friends or family members and only 13% of them thought about the idea by self.

Table No: 6
Motivational factors

Factors	No. of respondents	Percentage
Students	35	58
Financial rewards	20	33
Independence/ autonomy	35	58
Self-recognition	52	87
Passion	45	75
Personal growth	36	60
Happiness	30	50
Desire to do something	42	70
Encouragement	39	65
Promising demand for the product	32	53
To have fun	25	42
Family security	12	20

Source: primary data

Self-recognition, passion, desire to do something, encouragement and personal growth are the main motivating factors that influence the netpreneurs in north region of Kerala.

Table No: 7
Purpose of the Startup

Category	Number of respondents	Percentage
Students	35	58
To gain Self-confidence	32	54
Recognition	8	13
Fame	0	0
Self sufficiency	20	33

Source: primary data

The table no.7 shows that the 54% people started this business to gain self-confidence, 33% of them entered to the business for attaining self-sufficiency only 13% of them chose this business for recognition.

Table No: 8
Duration of Netpreneurship

Category	Number of respondents	Percentage
Below 6 months	18	30
1 year	12	20
More than 1 year	30	50

Source: primary data

Most of the people in this study are dealing with this business for more than a year. 50% respondents are started their business before 1 year.

Table No: 9
Degree of Interest towards business

Category	Number of responses	Percentage
High	46	77
Medium	14	23
Low	0	0

Source: primary data

The data shown in the table represents the interest rate of netpreneurs towards their business. From the total respondents, 77% of people have high

interest on the business, and remaining 23% of people has medium interest on their business.

Table No: 10
Mode of delivery of product

Category responses	Number of respondents	Percentage
In person	38	63.
Other courier services	22	37
Own courier services	0	0

(Source: primary data)

Around 63% of respondents are delivering directly to their customers. And the remaining 37% are using other courier services.

Table No: 11
Complaints resolution

Category respondents	Number of respondents	Percentage
Cash back	0	0
Product replacement	10	17
Both	4	7
Have got no complaints	46	76

Source: primary data

Majority of the respondents says they are not getting any complaints from the customers (76%), some of them (17%) are replacing the products as the resolution and few of them (7%) follows both cash back and product replacement.

Table No: 12
Promotion for the business

Category	Number of respondents	Percentage
Yes	50	83
No	10	17

Source: primary data

The table no 12 shows the percentage of people who are using multiple platforms for promotion of their business. 83% of the individuals are promoting their business through multiple online platforms and only 17% are not promoting through multiple platforms.

Table No: 13
Offers for customers

Category	Number of respondents	Percentage
Yes	52	87
No	8	13

Source: primary data

This data set in table shows the percentage of individuals who are giving offer/discounts to the customers. Most of the respondents are giving offers to their customers (87%). But only few of them (13%) are not giving any offers.

Table No: 14
Criteria for offer

Category respondents	Number of respondents (out of 52)	Percentage
Based on number of orders	36	69.23
Seasonal	2	3.85
Randomly	14	26.92

Source: primary data

This data in the table shows the percentage of individuals distributed to what type of criteria they are following for giving offers to the customers. Majority of the people are following traditional method, which is giving offers based on number of orders, and some people giving offers seasonally and few of them are facilitating offers to the customers randomly.

Table No: 15
Advertisement media used

Category	Number of respondents	Percentage
YouTube	6	10
Instagram	36	60
Facebook	12	20
Others	6	10

Source: primary data

The table no 15 shows the statistics of most preferred advertisement platform in Netpreneurship by individuals participated in this study. It is clear that the 60% of the individuals chosen Instagram and 20% of them selected Facebook. 6% each are the percentage of people who selected Facebook and other medias for ads.

Table No: 16
Income from the business

Category	Number of respondents	Percentage
2K – 5K	20	33
5K – 10K	14	24
More than 10K	26	43

Source: primary data

The table no 16 shows the statistics of income earned during an average of one month. 43% of respondents are earning more than 10000 during a month. And remaining 24% of respondents are earning between 5000-10000. Only 33% are earning in between 2000-5000.

Table No: 17
Mode of Payment

Category	Number of respondents	Percentage
Bank transfer	6	10
Cash	12	20
Both	42	70

Source: primary data

The data shown in the table no 17 explains the mode of payment chosen by the netpreneurs participated in this study. Majority of them provides both cash and bank transfer methods for payment, 20% of the prefer cash mode than bank transfer and only 10% provides the bank transfer mode as only option.

Testing Of Association Between Use Of Multiple Promotion Methods And Income From Business

Ho: There is no association between Use of multiple promotion methods and Income from business

H1: There is association between Use of multiple promotion methods and Income from business

Table No: 18

Use of multiple promotion methods	Income from business			Total
	2000-5000	5000-10000	Above 10000	
Yes	18	11	21	50
No	2	3	5	10
Total	20	14	26	60

$$\chi^2 = \sum \left(\frac{(O - E)^2}{E} \right)$$

Level of significance : 0.05

Critical region : $\chi^2 > \chi^2_{\alpha, df}$

$$\chi^2 > 5.991$$

$$\begin{aligned} \text{Degree of freedom} &= (r-1)(c-1) \\ &= (2-1)(3-1) \\ &= 2 \end{aligned}$$

$$\text{Test statistics } \chi^2 = \sum \left(\frac{(O - E)^2}{E} \right)$$

E= (Column total * Row total)/Number of observation

Table No: 19

CALCULATION OF CHI SQUARE VALUE

O	E	O-E	(O-E)- 1/2	((O-E)- 1/2) ²	((O-E)- 1/2) ² /E
18	16.7	1.3	0.65	0.4225	0.025
2	3.3	1.3	0.65	0.4225	0.128
11	11.7	0.7	0.35	0.1225	0.010
3	2.3	0.7	0.35	0.1225	0.053
21	21.7	0.7	0.35	0.1225	0.0056
5	4.3	0.7	0.35	0.1225	0.028
Total					0.2496

Comparison: $\chi^2 < \chi^2_{\alpha, df}$

$$0.2496 < 5.991$$

Conclusion: The computed value of χ^2 is less than the table value. So, we accept the null hypothesis. i.e., there is no association between use of multiple promotion methods and Income from business.

Table No: 20
Business tie up

Category responds	Number of responds	Percentage
Yes	4	7
No	56	93

Source: primary data

The table describe the tie up that the respondents have with other business. 93% of respondents do not have any tie up and remaining 7% have tie up with other business.

Table No: 21
Mode of collecting review

Category responds	Number of responds	Percentage
Social media	54	90
E mail	0	0
Conducting survey	6	10

Source: primary data

The table no 21 shows that 90% of respondents are using social media for getting response and around a 10% is using survey method

Table No: 22
Mode of expansion plan

Category responds	Number of responds	Percentage (out of 50)
Introduce new product	39	78
Tie up	0	0
Retail store	11	22

Source: primary data

The data in the table 22 shows details regarding mode of expansion plan in future.78% of respondents are planning for future expansion of their business by adding up new product line. While 22 %are interested to start a new retail store.

Table No: 23
Impediments Of Netpreneurs

Problems	Number of respondents	Percentage
Finance	20	33
Lack of knowledge	10	17
Marketing problem	25	42
Lack of self confidence	30	50
Child care/house work/studies	35	58
competition	42	70

New technology	45	75
Lack of support	13	22
Health issues	25	42
Time management	52	87
criticism	33	55
Finding customers	41	68
Payment issue from customers	29	48
Stress	17	28
Labour problem	6	10
Challenges in product delivery	18	30
Difficulty in collecting raw material	55	92

Source: primary data

Though majority of the netpreneurs under study are students and house wives they have a time management problem. The major proble faced by netpreneurs are difficulty in collecting raw Materials, competition from others, new technology.

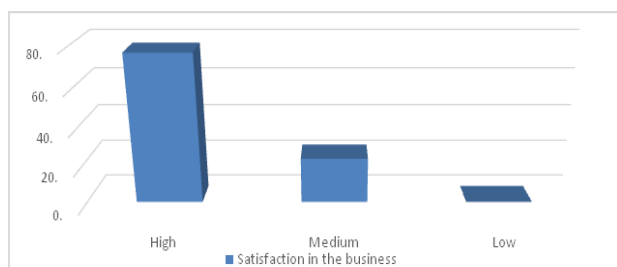
Table No: 24
Level of satisfaction

Category	Number of responses	Percentage
High	46	77
Medium	14	23
Low	0	0

Source: primary data

The data shown in the table 24 represents the satisfaction of netpreneurs towards their business. From the total respondents, 77% of people has highly satisfied on the business, and remaining 23% of people have medium level of satisfaction on their business.

Figure No 1
Satisfaction of Netpreneurs



Suggestions

1. Since the majority of the participants in the netpreneurship are women, large scale women empowerment can be done by conducting awareness programs to women who are currently not in to any business or career.
2. Employment opportunity is more in the area of raw materials supply since some of them faced difficulty in collecting raw materials. This will have a collective job opportunity in the whole process.
3. It is suggested that to introduce this kind of business awareness as an optional curriculum, so that students can get clarity about the opportunities in this area
4. Government schemes can be brought to the society for motivating this kind of business so that overall economic development can be planned.
5. The awareness of this business opportunity can be done in rural areas so that employment will be increased in the country
6. By providing rewards and recognition for the best performer in the area of this business will be a motivation and inspiration to all.

9. Conclusion

In this era netpreneurship has a wide scope because internet is becoming the major part in current generation. While promoting training and motivating netpreneurship the country can attain growth in employment and self-sufficiency. It gradually leads to development of the economy there by development of the country. This study is

entitled with the prospects and challenges faced by netpreneur in Kannur district. students and house wife engaged in netpreneurship for earning by their passion. It revealed that majority of women are entered in this kind of business for attain self-confidence and self-sufficiency.

They are facing many challenges such as difficulty while delivering the product, collecting payments, some kind of health issues, difficulty while collecting raw materials. Netpreneurship become highly motivating element in most of the individual because it does not need any major initial investment and this can be done from anywhere at any time.

10. REFERENCES

1. Ahmadpoor Dariani, Mahmoud, 2000. Entrepreneurship: Definitions, Theories and Models, Third edition, Tehran, Spring
2. Hoselitz, B.F., "Entrepreneurship and Economic Growth" The American Journal of Economics and Sociology, Vol.No.12, No.1 Oct 1952
3. Khanka, S.S., Entrepreneurial Development, S. Chand & Co, Ltd, New Delhi, 2004, pp. 1-2
4. Sheobahal Singh., Entrepreneurship and social change, Rawat Publications, Jaipur, 1995, p. 120
5. Subba Rao, P.D and Sundaram on Entrepreneurial Challenge in under developed sectors, Kanishka Publishers, New Delhi, 1993 p.9

Web sites

1. http://shodhganga.inflibnet.ac.in/bitstream/10603/64899/8/08_chapter1.pdf
2. <http://netpreneur.co.in/>
3. https://www.youtube.com/watch?v=LR-ZUjayd_w
4. <https://medium.com/@joyceud27/understanding-the-term-netpreneurship-1a5bd2eb7d0c>
5. <http://netpreneur.co.in/netpreneurship-internet-business-entrepreneurship-the-time-is-now-the-message-is-forever/>

Deep Learning Based Handwritten Text Recognition from Images

Aswathi Gopan

ICT Academy of Kerala
Trivandrum, India
aswathigopan1988@gmail.com

Riji N Das

ICT Academy of Kerala
Trivandrum, India
riji.n@ictkerala.org

***Abstract:** Handwritten Text Recognition (HTR) from images is a challenging task for machines. The variation in the writing styles of individuals makes HTR even difficult. Traditional HTR approaches extract language dependent features from images through various Computer Vision (CV) techniques. Natural Language Processing (NLP) powered by Convolutional Neural Network (CNN) helps to extract features from images independent of languages. This paper proposes a HTR model based on Deep Learning, which extracts handwritten texts from images independent of the language. The model is built with TensorFlow (TF) and produces better results in less processing steps compared to other models.*

***Keywords:** Handwritten Text Recognition (HTR), Computer Vision (CV), Natural Language Processing (NLP), Convolutional Neural Network (CNN), Deep Learning, TensorFlow (TF).*

1. INTRODUCTION

As a result of digitisation, a lot of information is available as images. Digitisation is an ongoing process since despite having numerous digital writing tools, many people still prefer the traditional approach of writing down notes using pen and paper. The extraction of text from these digitised notes can provide immense amounts of meaningful information. Formerly used machine learning techniques like Hidden Markov Models (HMM) [2] [3], Support Vector Machine (SVM) etc are used for the text extraction. The machine learning models have limited performance due to the low learning capacity and manual feature extraction. These limitations can be overcome with the help of deep learning.

Artificial Intelligence (AI) and Deep Learning (DL) are gaining more importance these days. HTR is one of the major research areas under deep learning. The major challenges in this area are the writing styles of different individuals as well as the amount of characters. In order to develop a high-performance neural network, the model should be trained with a large dataset.

In this paper we use Convolutional Neural Network (CNN), which is one of the deep learning algorithms mainly used for image processing. One of the major advantages of CNN compared to other classification algorithms is that it requires less pre-processing steps [14][15]. Due to its particular structure, it can be applied to most of the computer vision problems [4]. CNN will process the input image and classify it to certain categories. In this work we break down the images into small patches and feed them into CNN which is then trained using the softmax function. One can also use CNN with CTC (Connectionist Temporal Classification) loss function for training, in that case there is no need for pre-segmentation.

II. RELATED WORKS

This section discusses some of the works done by researchers in the field of HTR.

A system was developed for the character reorganisation using the Fuzzy logic with the help of hamming neural networks [1]. The system was basically a VLSI structure immune to noise and with low power consumption.

A novel method for the word recognition system was proposed using a recognizer based on HMM [5]. The model in this system was based on decision tree clustering. Researchers basically experimented on Latin and Arabic languages using this system.

Another model was developed which detects the handwritten characters which are in English language without feature extraction. In this case the resized characters are directly used for the training purposes in such a way that all the pixels are features for training the neural network [6].

One of the researchers have developed a model in order to predict the handwriting sequence. The model makes use of a Self-organizing Map (SOM) for the feature extraction and Recurrent Neural Network as the learning module [3] [7]. The researchers use Japanese characters for the training and model evaluation.

A text recognition system was developed which converts the text from a hardcopy to the prescribed format [8]. They actually proposed a system where pre-processing and segmentation methods are used to detect a line. The system works good for almost all kinds of bimodal images.

During the past few years, CNN models are showing high accuracies for the character recognition systems [16]. A novel model was developed which made use of three networks: LeNet, DropConnect and Network3 along with their ensembles.[9]. The system analysed different transformations on the MNIST dataset.

Another previous work includes a system developed to distinguish the handwritten and the printed text from a document. They use the most challenging Arabic script for word recognition [11-13]. The classification model was developed using Expectation Maximization (EM) based neural network.

III.SYSTEM DESCRIPTION

In this section the steps followed in the implementation of the system is discussed in detail. The dataset used in this work is mainly from the IAM Handwriting database which is the biggest

database of English handwriting images. It consists of scanned texts [Fig.1] of more than 600 writers. With the help of this huge dataset we can train the model for better performance.

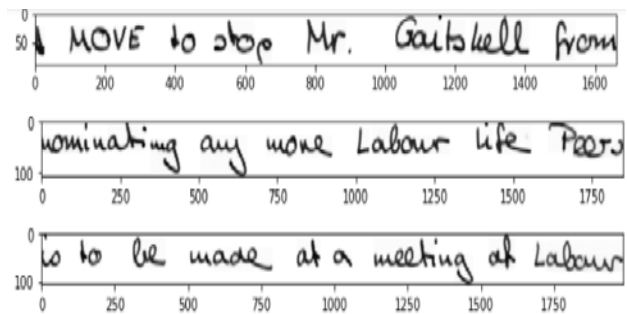


Fig.1. Sample images from IAM dataset

A. Normalization

During the training and evaluation of the model, we have to provide the images to the neural network. Several techniques can be used to scale down the pixel values. In this work we used the label encoder method. By doing so the pixel values are rescaled to 0-1 which is usually preferred by the neural networks. This type of scaling is usually referred to as normalization.

B. Data Splitting

Before proceeding further, we have to split our IAM dataset, which is now a normalized one, into three parts. One for training the model, the next one for testing the model and the last part for validation. The ratio we used for data splitting is 4:1:1.

C. Generating patches of data

The dataset which we have taken from IAM is now normalized and split to train, test and validation. The next task is to generate small patches from the data. This helps to send data as small patches of size 113 x 113 instead of passing sentences directly to the model. This can be achieved with the help of a generator function which is moved over each

sentence and will generate random patches of prescribed size. From the generated patches, we keep only 10 % of the random crops. One of the major advantages of using the patches is to make the model language independent.

The generator function is then applied to train, test and validation datasets.

D. Model Creation

As discussed earlier, CNN, the most common deep learning algorithm for image processing, is used for model creation. CNN uses two major operations namely, convolution and pooling for the feature extraction and these features are used for the classification. The basic architecture of a CNN is given in Fig.2.

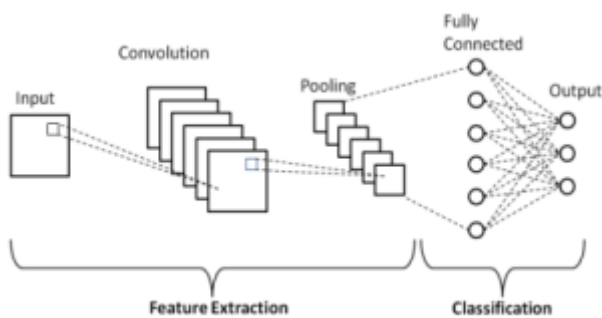


Fig.2. Basic CNN Architecture [10]

The above figure shows the simple CNN architecture with five layers which is grouped into two: feature extraction and classification. The feature extraction consists of input, convolution and pooling layers. The last two layers constitute the classification part which includes the fully connected and output layer.

- 1) Input Layer: This input layer provides the dataset to the convolution layer. The input in our work is the random patches generated from the normalised IAM data.
- 2) Convolution Layer: The role of convolution network is to reduce the input size by retaining all the features which are essential for the better

performance of a model. Each convolution layer consists of filters which move over the image from top to bottom depending upon the kernel size. We have to also specify the number of pixel shifts over the input with the help of the parameter stride. If the input image is not perfectly fitting with the filters, we have to use padding technique. With the help of padding, the feature map generated will be of the same size as that of the original image. The feature map generated is then passed to an activation function which determines the presence of features in a particular location of the image. The activation function used in our system is ReLu.

- 3) Pooling Layer: It is actually a dimensionality reduction technique which retains all the relevant information in the images. Pooling layer is used when we deal with a large dataset. Mainly three types of pooling techniques are there and they are Max pooling, Average pooling and Sum pooling. Max pooling method is used in our system.
- 4) Fully Connected Layer: This layer will flatten the image to a vector representation. This flattened data is then applied to the feed forward network during each iteration. After training the model along a series of epochs, it will be able to distinguish between the features.
- 5) Output Layer: This layer provides the final class of the input with the help of some activation functions. The activation function used here is the softmax function.

Normally a CNN network consists of multiple convolution layers followed by pooling. This helps to extract maximum features from the image which will enhance the model performance. In our system, we created a CNN model with three convolution layers with the activation function ReLu, and each layer followed by pooling. Finally, the output layer is designed using the softmax function and the model is optimised using Adam optimizer. The model summary is depicted in Fig.3.

Layer (type)	Output Shape	Param #
zero_padding2d_1 (zeroPaddin	(None, 115, 115, 1)	0
lambda_1 (Lambda)	(None, 56, 56, 1)	0
conv1 (Conv2D)	(None, 28, 28, 32)	832
activation_1 (Activation)	(None, 28, 28, 32)	0
pool1 (MaxPooling2D)	(None, 14, 14, 32)	0
conv2 (Conv2D)	(None, 14, 14, 64)	18496
activation_2 (Activation)	(None, 14, 14, 64)	0
pool2 (MaxPooling2D)	(None, 7, 7, 64)	0
conv3 (Conv2D)	(None, 7, 7, 128)	73856
activation_3 (Activation)	(None, 7, 7, 128)	0
pool3 (MaxPooling2D)	(None, 3, 3, 128)	0
flatten_1 (Flattern)	(None, 1152)	0
dropout_1 (Dropout)	(None, 1152)	0
dense1 (Dense)	(None, 128)	147584
activation_4 (Activation)	(None, 128)	0
dropout_2 (Dropout)	(None, 128)	0
dense2 (Dense)	(None, 64)	8256
activation_5 (Activation)	(None, 64)	0
dropout_3 (Dropout)	(None, 64)	0
output (Dense)	(None, 50)	3250
activation_6 (Activation)	(None, 50)	0
Total params: 252, 274 Trainable params: 252, 274 Non - Trainable params : 0		

None

Fig.3. Model Summary

E. Model Training and Performance Evaluation

The model created is trained using the train data which is already generated from the input dataset. While training the model, we have to specify the number of epochs and the batch size as well. After the successful training, the model is evaluated and we are able to attain 98% accuracy with the use of four epochs. Now our model is ready to extract the handwritten text from any image.

IV. CONCLUSION AND FUTURE SCOPE

A deep learning model based on CNN is proposed to automatically detect the handwritten

text from the given set of images. The model helps in the offline detection of the handwritten text from images through a language independent approach. This approach delivers high performance using CNN models with less processing steps compared to other models. There is still scope for improvement in this model. The model can be extended to detect a) the degraded text, b) digits, c) half characters, d) special characters and e) compound characters present in the images.

References

1. Wei Lu, Zhijian Li, Bingxue Shi . "Handwritten

Digits Recognition with Neural Networks and Fuzzy Logic" in IEEE International Conference on Neural Networks, 1995. Proceedings.

2. H. Hermansky, D. P. W. Ellis, and S. Sharma, "Tandem connectionist feature extraction for conventional HMM systems," International Conference on Acoustics Speech and Signal Processing, vol. 3, no. 28149, pp. 1635–1638, 2000.
3. Y. Yamashita and J. Tani, Emergence of Functional Hierarchy in a Multiple Timescales Recurrent Neural Network Model: A Humanoid Robot Experiment, PLoS Computational Biology, Vol. 4, e1000220, 2008.
4. Y. Le Cun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," International Symposium on Circuits and Systems, pp. 253–256, May 2010.
5. A.-L. Bianne, F. Menasri, R. Al-Hajj, C. Mokbel, C. Kermorvant, and L. Likforman-Sulem, "Dynamic and Contextual Information in HMM modeling for Handwriting Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 10, pp. 2066 – 2080, 2011.
6. J. Pradeep, E. Srinivasan, S. Himavathi. "Neural Network based Handwritten Character Recognition system without feature extraction" in International Conference on Computer, Communication and Electrical Technology ICCET 2011.
7. Shun Nishide, Hiroshi G. Okuno, Tetsuya Ogata, Jun Tani. "Handwriting Prediction Based Character Recognition using Recurrent Neural Network" in 2011 IEEE International Conference on Systems, Man, and Cybernetics.
8. Tridib Chakraborty, Chwdhury Md Mizan, Suparna Karmakar. "Text recognition using image processing", International Journal of Advanced Research in Computer Science, 8 (5), May-June 2017, 765-768.
9. Siham Tabik, Daniel Peralta, Andrs Herrera-Poyatos, Francisco Herrera. "A snapshot of image Pre-Processing for convolutional neural networks: Case study of MNIST" in International Journal of Computational Intelligence Systems 10(1):555 January 2017.
10. Phung, V.H.; Rhee, E.J. A Deep Learning Approach for Classification of Cloud Image Patches on Small Datasets. J. Inf. Commun. Conver. Eng. 2018, 16, 173–178, doi:10.6109/jicce.2018.16.3.173.
11. F. Farooq, K. Sridharan, and V. Govindaraju, Identifying Handwritten Text in Mixed Documents, ICPR 2006, 18th International Conference on Pattern Recognition, v. 2, pp. 1142 – 1145, 2006.
12. S. Violante, R. Smith, and M. Reiss, A Computationally Efficient Technique for Discriminating Between Hand-Written and Printed Text, IEEE Colloquium on Document Image Processing and Multimedia Environments, 2 Nov., pp. 17/1 – 17/7, 1995.
13. J. Koyama, M. Kato, and A. Hirose, Local-spectrum-based distinction between handwritten and machine-printed characters, 15th IEEE International Conference on Image Processing, 12-15 Oct., pp. 1021 – 1024, 2008.
14. S. Imade, S. Tatsuta, and T. Wada, Segmentation and Classification for Mixed Text/Image Documents Using Neural Network, Proceedings of the Second International Conference on Document Analysis and Recognition, 20-22 Oct., pp. 930 – 934, 1993.
15. B. V. S. Murthy. "Handwriting Recognition Using Supervised Neural Networks" in the International Joint Conference on Neural Networks, 1999. IJCNN '99.
16. Shao Y., Wang C., Xiao B., Zhang Y., Zhang L., Ma L. (2010) Text Detection in Natural Images Based on Character Classification. In: Qiu G., Lam K.M., Kiya H., Xue XY., Kuo CC.J., Lew M.S. (Eds) Advances in Multimedia Information Processing – PCM 2010. PCM 2010. Lecture Notes in Computer Science, vol 6298. Springer, Berlin, Heidelberg.

COVID-19 Analysis and Outbreak Forecast using Machine Learning Techniques

Aswathi Gopan

Dept. of Data Science & Analytics
ICT Academy of Kerala
Technopark, Trivandrum
aswathigopan1988@gmail.com

Ijaz Ahammed

Dept. of Data Science & Analytics
ICT Academy of Kerala
Technopark, Trivandrum
ijazahmed.ep@gmail.com

Abstract : *The outbreak of COVID-19 has brought a real threat to the living society. The whole world is fighting hard to get rid of this pandemic. In this paper we present analysis of COVID -19 transmission and predicts its spread and recovery rate using suitable machine learning (ML) models. The analysis and prediction are done both worldwide and in India. We use both regression models such as Linear Regression, Support Vector Machine (SVM) and time series forecast models such as Prophet, AR, ARIMA for the prediction.*

Keywords: COVID-19, ML, SVM, Prophet, AR, ARIMA, RMSE

1. INTRODUCTION

Coronaviruses are one among the viruses that cause severe respiratory problems in human beings. The outbreak of coronavirus was reported from Wuhan, China by December 2019. It spread quickly throughout the world. Thousands of new cases are reporting from various countries and the mortality rate is also increasing day by day. The virus spreads mainly through the direct contact and the major challenge facing now is that new cases reporting today does not show any symptoms. Most of the countries declared lockdown in order to curtail the spread of this severe pandemic. Precise forecasting of the spread of COVID-19 will help various countries to take precautions against the virus spread.

We propose traditional machine learning models to forecast the virus spread [1],[2],[3]. Machine learning is a prominent field used now a days to solve real life problems. Machine learning may be defined as an application of AI which uses the data to learn and improve the performances. Forecasting is one among the significant areas of ML. This study aims to create a precise model to forecast the spread of COVID-19.

Regression is one of the traditional supervised ML technique which can be used for forecasting purposes. In this paper we use Linear Regression and Support Vector Machine (SVM) for the forecasting. These traditional models are compared with some time series models like prophet, AR and ARIMA models. Root mean square error (RMSE) value is used for the performance evaluation of various models and comparing RMSE values. ARIMA model performed well. Therefore, we prefer ARIMA model to forecast the new cases and recovery rate throughout the world and in India.

The paper is organized as follows. Section II briefly explains the methodology including data set description, data processing steps, exploratory data analysis, various ML models. Section III includes results in detail and finally conclusion in section IV.

I. METHODOLOGY

This section briefly explains the proposed ML models. The data set which is used for this paper has been obtained from Kaggle. After performing preprocessing steps some analysis is done on the data and after that we performed some ML models to forecast the spread of COVID-19.

A. Data Preprocessing

The data set which is used in this project is a real time data which includes the confirmed, deaths and recovered cases in various countries date wise [8],[11]. The data available in this data set is from Jan 22nd 2020 to July 4th 2020. The preprocessing steps includes

- ▶ The data set is properly cleaned by dropping the columns that have the null values.
- ▶ For the further analysis purpose, we add some additional columns including "Growth Factor", "Recovery rate" and "Mortality rate".

B. Exploratory Data Analysis

The exploratory data analysis is done on the given dataset as three parts: worldwide, country wise and the impact of COVID-19 in India [9],[10].

1) World wide Analysis: In order to perform world wide analysis the dataset is grouped according to the data wise and now we have 164 samples.

a) Confirmed / Death cases over time: Fig.1 depicts the confirmed/death cases over time throughout the world. It's clear from the figure that the confirmed cases are increasing exponentially and the death cases are also increasing day by day. From this it's clear that the situation is going worse day by day and proper awareness should be given to the people regarding this concept. It's also evident that how fast this epidemic is spreading all over the world. Since no vaccine has been discovered the only way to reduce its impact is social distancing and proper hygiene.

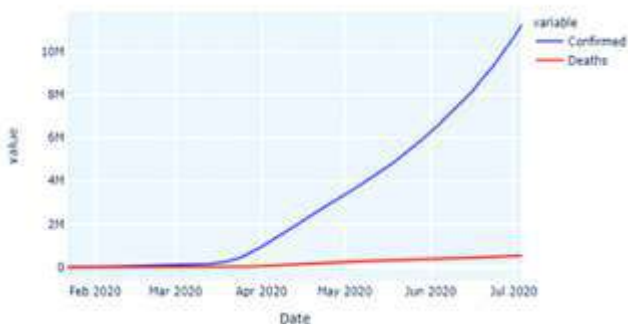


Fig.1 Worldwide confirmed/ death cases

b) Recovered / Active cases over time: Fig.2 depicts the recovered / active cases throughout the world. We can see that during the month of May the active cases are high compared to the recovery cases. But now the recovery cases are increasing which is a good sign.

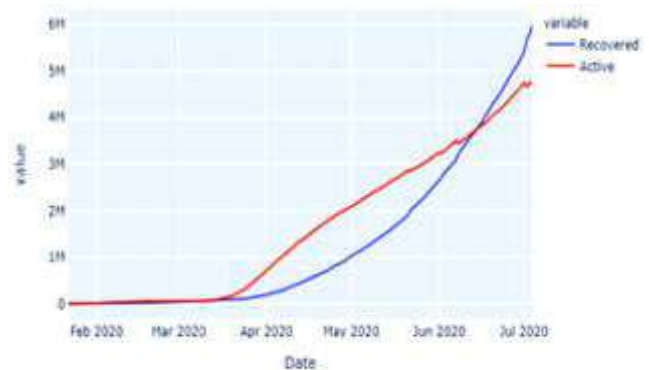


Fig.2 Worldwide recovered/ active cases

c) Recovery rate / Mortality rate over time: The recovery rate and mortality rate are two important factors which determine how effectively we are overcoming this epidemic. From Fig.3 it's evident that the recovery rate is increasing and at the same time there is a slight decrease in the mortality rate.

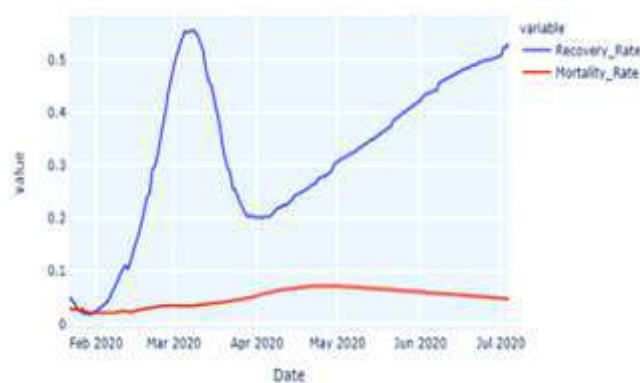


Fig.3 Worldwide recovery/ mortality rate

2) Country wise Analysis: Before going to the country wise analysis we consolidated the dataset till July 4th 2020. The dataset consists of the data of 187 countries and the analysis is done accordingly.

- a) **Confirmed Cases Trend:** Fig.4 depicts the trend in the confirmed cases of top 10 countries. US is having more confirmed cases followed by Brazil.

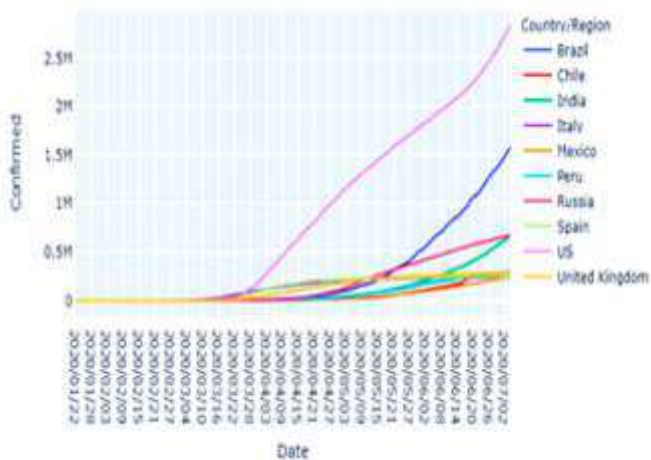


Fig.4 Confirmed cases trend of top ten countries

- b) **Mortality trend:** Fig.5 shows the mortality trend of top 10 countries. The increase in mortality rate is not a good sign. Again US is having high mortality rate among 187 countries.

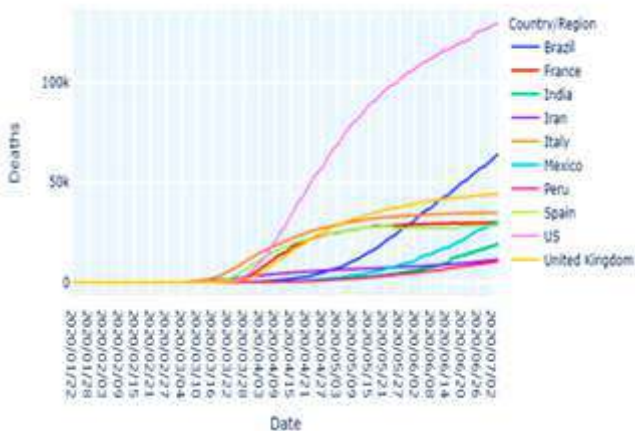


Fig.5 Mortality trend of top ten countries

- c) **Recovery rate :** For the better control over the spread of COVID-19 the recovery rate should be high for each country with minimum mortality rate.

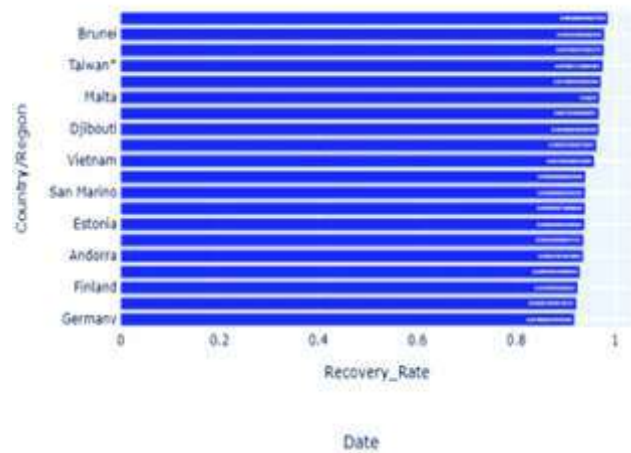


Fig.6 Recovery rate of top ten countries

Fig.6 shows the recovery rate of top ten countries. Brunei have the highest recovery rate among 187 countries under analysis.

- 3) **Analysis of COVID-19 in India:** Similar to what we have done in the world wise analysis, a small EDA is done on the cases in India.

- a) **Confirmed / Death cases over time :** Fig.7 illustrates the confirmed/death cases in India from January 2020 to July 2020. Till March the confirmed cases in India was too less, but after that it increases exponentially.

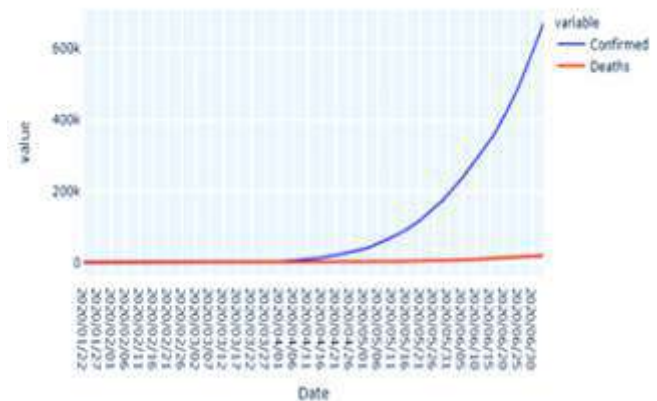


Fig.7 Confirmed/ death cases in India

- b) **Recovered / Active cases over time:** The active cases in India is high during the months April and May compared to the recovery rate. After May, there is an improvement in the recovery rate which indicates the people are able to overcome this situation. Fig.8 illustrates the same.

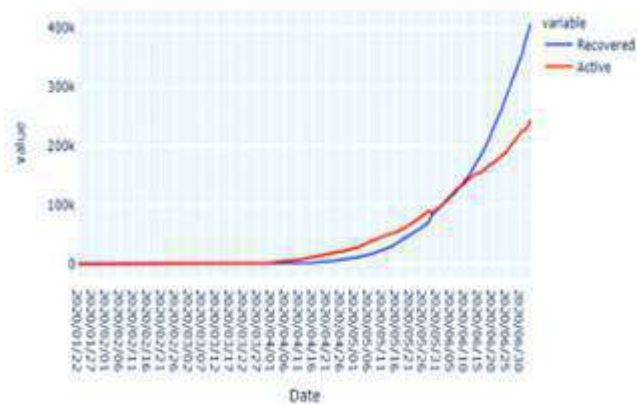


Fig.8 Recovered/ active cases in India

- c) New Cases over time: In India, first COVID-19 case reported in Kerala on January 30th 2020, when a student returned home from Wuhan University. Second and third cases are again reported from Kerala where the virus is detected in two more students from Wuhan university. After that the new cases increase slowly. The government implemented lockdowns in order to curtail the virus spread. By the end of June, India fourth worst-hit country in terms of COVID-19 cases and the cases are approaching to 1 million. Fig.9 depicts the trend in the new cases over India from January 2020 onwards.

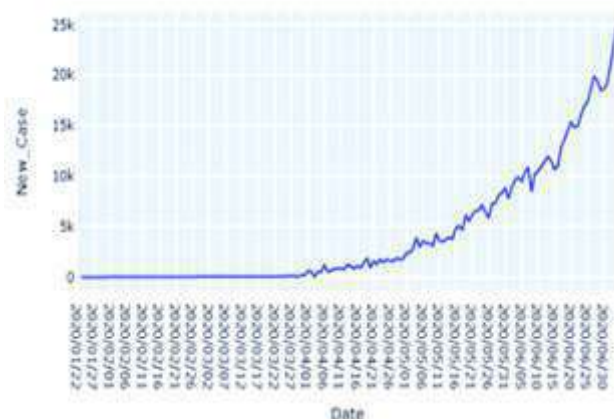


Fig.9 New cases in India

- d) Recover rate / Mortality rate over time: Fig.10 depicts the recovery/mortality rate in India. The recovery rate drops drastically during the month of April. Afterward, the recovery rate is increasing and the mortality rate is also not so high.

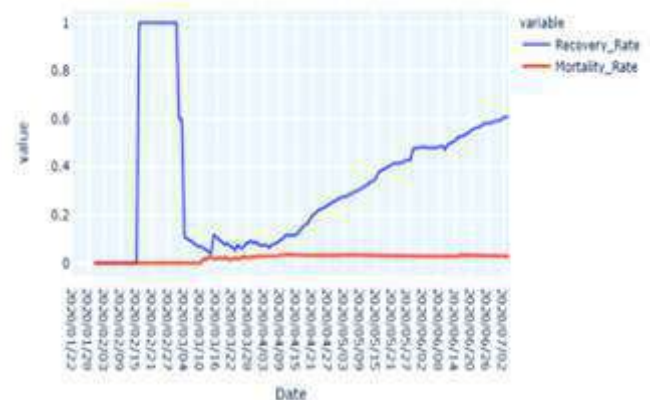


Fig.10 Recovery/ mortality rate in India

- e. Machine Learning Models

Supervised machine learning models are used to forecast COVID-19 outbreak [4],[6]. The dataset is divided into train and test with test size as 0.2. The trained models are then made to predict the result for the test data. In our project the models are trained to predict the confirmed cases. Models are evaluated using the performance parameter, RMSE value.

- 1) **Linear Regression:** It is a statistical approach which will find the relationship between the dependant and independent features. It is more attractive model for the predictive purpose, since it's too simple. The relation between the dependant (y) and independent (x) features is shown in Eq.(1).

$$b_0 + b_1 * x = y \quad (1)$$

Here b_0 represents the intercept and the b_1 represents the slope. This equation can be used to predict the value of target variables based on the predictor variables. The primary goal of the ML algorithm is to find the best fit values of b_0 and b_1 from the train data provided. With the help of these values, model will try to get a best fit regression line.

The best fit implies the minimum difference between the actual and predicted values.

2) **Support Vector Machine:** One of the most common supervised machine learning model used for both regression and classification problems is support vector machine(SVM) [5],[7].Inorder to transform the input data to the desired form, special functions called kernels are used. In case of regression problems, a hyperplane and a decision boundary is considered at a certain distance from the hyperplane. Then we are trying to make all the data points within this decision boundary.

In case of simple regression, we are trying to make the error minimum. While in case of support vector regression, we are trying to make the error value within some threshold value[12].

3. **Prophet:** Prophet is an open source software from Facebook and is mainly used for forecasting purposes. This can be used to forecast time series data. The dataframe applied to this model should contains only two columns. One is for timestamp and other for the numeric data. In our project, one column consists the data and other one the confirmed cases. Fig.11 shows the prediction using the prophet model using the trained model.

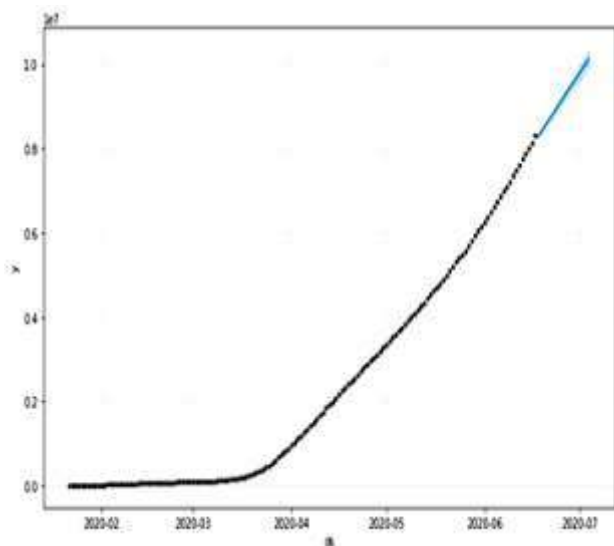


Fig.11 Prediction of confirmed cases using prophet

4) **Auto Regressive Model:** It is another ML model that accurately forecast the time series data.It is

somewhat similar to normal regression.As an input to the regression equation, the data points from the previous times steps are applied to predict the value for the next time stamp.The number of preceeding values used in the regression equations for forecasting is defined as an “order” of the model.Its called autoregression because it uses the previous times stamp of the same input variable to predict the future values. In our model,we use the order vale as one, first order autoregression.Fig.12 shows the actual and predicted values obtained using AutoRegressive(AR) model.The blue line indicates the actual values and the red line indicates the predicted values.

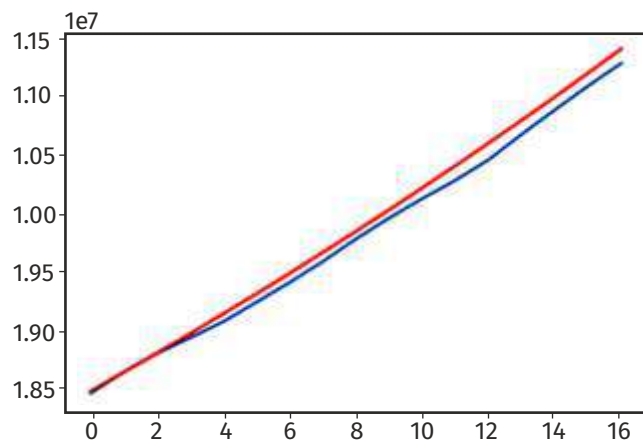


Fig.12 Prediction of confirmed cases using AR model

1) **AutoRegressive Integrated Moving Average (ARIMA)** : It is another ML time series model mainly used for time series forecasting.AR-denotes the Auto Regression indicating the prediction using lagged values,I- Integrating which is used to make the time series stationary and MA-Moving Average. In order to create the ARIMA model, we have to specify three parameters (p,d,q).

- ♦ p - number of lag observations
- ♦ d - number of differencing to be performed
- ♦ q – moving average order

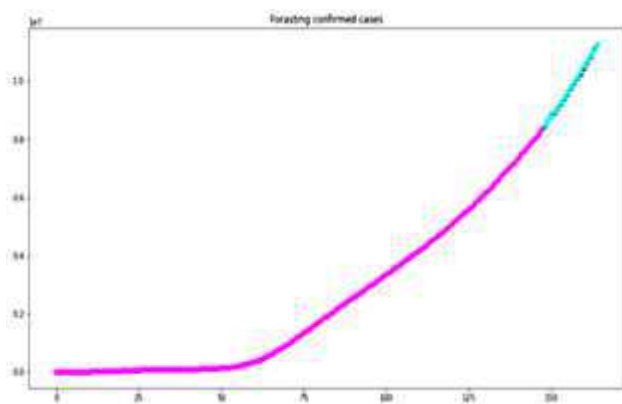


Fig.13 Prediction of confirmed cases using ARIMA model

III. RESULTS AND DISCUSSION

As discussed earlier all the ML models are trained and tested to predict the confirmed cases. The performance of the models is evaluated with the help of RMSE values. The below table shows the RMSE value of each model. From the table it's clear that the ARIMA model is having low RMSE value and because of that we use this ARIMA model to predict the outbreak of COVID-19 throughout the world as well as in India. Along with that the same model is also used to forecast the recovery rate also. The forecasting is done for the time stamp July 5th to July 31st 2020. From the forecasted data, it's evident that this infectious virus is spreading fast throughout the world.

TABLE 1. Model Performances on Forecasting Confirmed Cases

Model	RMSE Value
Linear Regression	86433.55
SVM	89146.25
Prophet	657540.91
AR	87218.63
ARIMA	31017.92

A. Forecast Using ARIMA Model

As discussed earlier, the new cases and the recovery rate are forecasted using ARIMA models for the time stamp from July 5th to July 31st 2020.

1) **Forecasting New Cases** : The new cases are forecasted for both worldwide and in India. By examining the forecasting in Fig 14 and Fig.15, it's evident that the cases are increasing day by day. So it's very important that all should give at most care to reduce the virus spread.

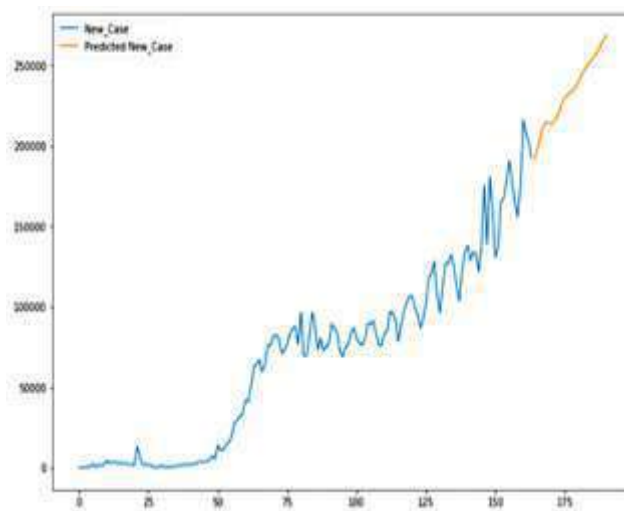


Fig.14 Forecasted New Cases in World

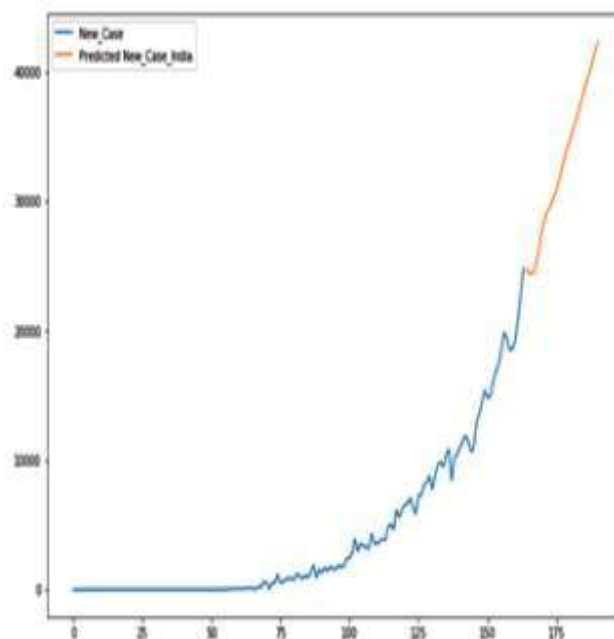


Fig.15 Forecasted New Cases in India

1) **Forecasted Recovery Rate** : The ARIMA model is also used to forecast the recovery rate throughout the world and in India. Even though the new cases are increasing day by day, there is a slight relief regarding the recovery rate .Fig 15 and Fig.16 shows the forecasted recovery rates.

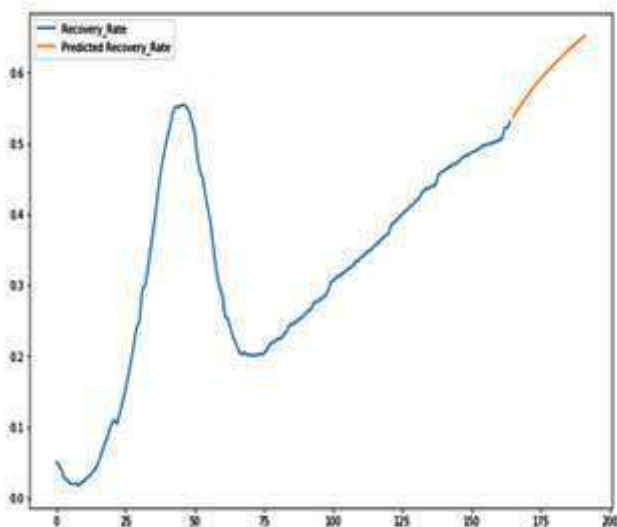


Fig.16 Forecasted Recovery Rate in World

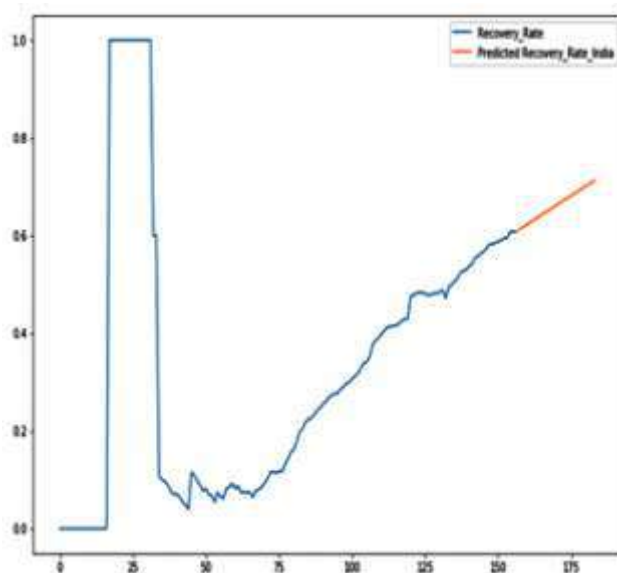


Fig.16 Forecasted Recovery Rate in India

IV. CONCLUSION

In this paper, we had done the virus spread forecasting of COVID-19 virus with the data set collected from Kaggle. Most popular ML forecasting models were implemented and was able to achieve good results with the help of ARIMA model. The forecast using ARIMA model will be a great help for the authorities to predicting the spread of epidemic and to take necessary timely decisions. In future we plan to build model using neural networks with updated dataset.

Acknowledgment

This work was done as a part of Certified Specialist Course in Data Science and Analytics from ICT Academy, Kerala. We would like to extend our sincere gratitude to all the members of ICT Academy for their valuable comments and timely support which helped us for the successful completion of this project.

References

1. Andrea Apolloni, Chiara Poletto, and Vittoria Colizza. "Age-specific contacts and travel patterns in the spatial spread of 2009 h1n1 influenza Pandemic", BMC infectious diseases, 13(1):176, 2013.
2. Dhiraj Dahiwade, Gajanan Patle, and Ektaa Meshram, "Designing disease prediction model using machine learning approach", In 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), pages 1211–1215. IEEE, 2019.
3. Anjan Nikhil Repaka, Sai Deepak Ravikanti, and Ramya G Franklin. "Design and implementing heart disease prediction using naives bayesian", In 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), pages 292–297. IEEE, 2019.
4. Olivera Stojanović, Johannes Leugering, Gordon Pipa, St'ephane Ghazzi, and Alexander Ullrich. A bayesian monte carlo approach for predicting the spread of infectious diseases. PloS one, 14(12), 2019.

5. S.A. Shreyas, S Abishek , N. Radhika, "Performance Analysis of Linear Grid Stability Using Classifiers and Advanced Ensemble Techniques", Journal of Advance Research in Dynamical & Control Systems, Vol. 11, 03-Special Issue, 2019
6. S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and machine learning forecasting methods: Concerns and ways forward," PLoS ONE, vol. 13, no. 3, Mar. 2018, Art. no. e0194889
7. C.-H. Wu, J.-M. Ho, D.-T. Lee Travel-time prediction with support vector regression IEEE Trans Intell Transp Syst, 5(4), pp. 276-281, 2004.
8. Johns Hopkins University Data Repository. Cssegis and data. Accessed: Mar. 27, 2020. [Online]. Available: <https://github.com/CSSEGISandData>.
9. F. Rustam, I. Ashraf, A. Mehmood, S. Ullah, and G. Choi, "Tweets classification on the base of sentiments for US airline companies," Entropy, vol. 21, no. 11, p. 1078, Nov. 2019.
10. Y. Grushka-Cockayne and V. R. R. Jose, "Combining prediction intervals in the m4 competition," Int. J. Forecasting, vol. 36, no. 1, pp. 178-185, Jan. 2020.
11. <https://github.com/CSSEGISandData/COVID-19>, Last accessed: 7 st May 2020.
12. Raj, J. S., & Ananthi, J. V, "Recureent Nueral Networks and Non Linear Prediction In Support Vector Machines", Journal of Soft Computing Paradigm (JSCP), 1(01), 33-40, 2019.



Information & Communication Technology Academy of Kerala

(A Govt. of India supported, Govt. of Kerala partnered Social Enterprise)

L-9 Thejaswini, Technopark Campus, Thiruvananthapuram, Kerala, India - Pin: 695 581
Ph: +91 471 2700811 / Fax: +91 471 2700812 / e-mail: info@ictkerala.org / www.ictkerala.org

Follow us at www.facebook.com/ictkerala