Appealing AI in Appalling Covid-19 Crisis and the Impending

Debanjan Banerjee, Anindita Banerji, and Arunita Banerji

Abstract The outbreak of the deadly Covid-19 virus has snatched smiles from everyone's face and now the entire world has been affected directly or indirectly by the effects of the virus, this virus keeps on mutating due to which there is no proper medicine or a final vaccine that assures it will curb the spread of the virus, major countries all over the world has lost more people than in a war and is still losing its people even after getting fully vaccinated. The horror is so much imbibed in each human it seems unrealistic to even think that the world will be normal ever again. This outbreak of the unknown virus is certainly a black-swan event that has annihilated people economically, emotionally, and socially and has made each individual realize the importance of one's health and how to be a responsible person by taking care of whatever finances one has, as in unprecedented times savings are the only resort left with a person. It is a testing time and everyone is at war, we all are soldiers in this pandemic and our health care workers, administration, and government are trying their best to stop the spread of the disease as it has killed more than four lakh people in India only and in the world tally is more than forty lakhs with numbers increasing. In this appalling situation when everything has been shifted to online mode solutions must be looked at in more technologically driven methods, in today's world due to rapid advancement in the IT and computer science sector there are ways to track the next rising hotspot of the virus and how it can be contained by taking swift actions if predicted within a particular time frame. Data collection, data analysis, and studying trends can help in assessing the upcoming threats, and in this manner, new job opportunities can also be created as it will involve people being prepared with limited medical knowledge to cure the people affected with the virus. In these times government and administration must adopt technologically backed solutions

D. Banerjee (✉)
Kalinga Institute of Industrial Technology (Deemed to be University), Bhubaneswar, India
e-mail: debanjan.banerjee98@gmail.com

A. Banerji
Tata Institute of Social Sciences, Guwahati, India

A. Banerji
FORE School of Management, New Delhi, India

S. S. Rautaray et al. (eds.), *Data Science in Societal Applications*, Studies in Big Data 114, https://doi.org/10.1007/978-981-19-5154-1_2

which will help the system to make accurate decisions based on real-time data-driven modeling capable of identifying the relevant information.

1 Introduction

This chapter will explore how AI can help in tackling pandemic-like situations with prior analysis by collecting relevant data, identifying the underlying patterns of influence and giving the most appropriate results after analyzing the suitable solutions with reference to the present dynamics of events along with the historical trend relevance. AI can be used very diligently to identify the virus hotspot areas, what are the trends in recent times, and how they can be tackled including various other factors like how people are following the protocols and how people are getting vaccinated, whether vaccination is helping in bringing down the fatality rate or chances of spread of the virus is decreasing; this all can be estimated with a set of models which work on relevant data which uses past events to give predictions using probability functions induced with complex mathematics.

It is not a hidden fact that the second wave of Covid-19 ripped India apart as the administration was not ready nor there were proper protocols followed plus a crisis in health care units and hospitals was seen, this was due to unpreparedness and too much relaxation of the protocols. In such a situation AI comes into play as it can analyze tons of data and predict what can arise in the future. The predictions based on recent dynamics can give results that are not too far away from the truth, with prior information and tracking of people the administration can take required steps to control the situation and the health care system can be alerted so that they can be prepared with enough resources. In this chapter, various tools and models are discussed which can be used for future predictions using machine learning, statistical models, and complex mathematical programs. The end goal of the proposed framework is to build an AI-powered Autonomous Assistance Algorithm which will proactively not only help the governing authorities in tackling pandemics like Covid-19 but will be playing a key role in assisting humanity to fight back to normal using its intelligence.

2 Covid-19: The Key Observations and Learnings

In this unfortunate and unprecedented situation, the entire world has faced immense disruption and humongous loss of lives. Not a single country was prepared for such a situation where everything seems out of control and the entire administration and health care system fell apart, ripping the countries to their fullest extent. The virus is attacking in waves and is mutating thus, no one is sure about the third wave, many experts say that it is inevitable but its spread can be curbed if required measures are taken on time and the system is prepared for it.

In the past two waves the key observations and learnings are as follows:

- Administration Challenges:
  The entire situation was handled based on how other countries were handling it, India's huge population and the existing nature of work life was not ready for such a shock, the government ministries failed to detect the source of the virus and were not able to identify the viable hotspots before the cases in those particular areas reached its peak. The unpreparedness of the administration can be seen in every state, the second wave was more catastrophic and it could have been like this if the administration had taken enough measures to create more healthcare facilities and had banned all the gatherings for and year after understanding the sources and hotspots of the virus. Due to a lack of efficient management the doctors were also not provided with the advanced requirements that were advised, they had to provide their services with limited resources. There were no mandatory checks at the airport and the incoming passengers from countries were not entitled to mandatory quarantine for fourteen days as well, this all exposed the residents of that state exposed to the virus in an unimaginable manner. Once the situation started worsening then only lockdown and bans were issued, the virus cannot be stopped from spreading by taking spontaneous actions it can be stopped by imposing precautionary measures in its toughest forms.
- Unemployment:
  One of the adversely affected is the middle-class youth, who had to undergo immense pressure and have the responsibility of their families. In this situation, many young adults lost their family and friends hence making them the sole breadwinner and this compelled young students to quit education and start searching some jobs to support their families but not many were successful in their endeavors, the world went through mass-layoffs and may employees who were in heavy machinery, etc., lost their jobs and had a family to look after, they were compelled to do petty jobs. Due to a sudden halt in all the day-to-day work, many migrant laborers lost their jobs, had no source of income, faced severe food shortages, and were uncertain about the future. Government had made provisions for them, but those provisions did not reach them in a systematic manner. These workers then wanted to go back to their villages as in several rural areas the spread of the virus was less and farming along with other livelihood work were still going on. The crisis reached its peak when migrant workers were trying to go back to their homes which created a lot of crowding in bus stations and railway stations.
- Impact on Academia:
  In the wake of the Covid-19 pandemic, the entire education system has shifted to the online mode which has exposed the students, teachers, and educators to a pool of challenges. Many students do not have the proper resources or live far off to catch good signal strength, it has become immensely tiring for the students to attend lectures online and study at home all by themselves with some help from their parents. There is no personal attention toward the weaker students and in rural parts, the online mode of education still has a long way to go to be reachable to the students, during this pandemic 11% of the rural households have

bought smartphones but cannot access many things. The students who have all the facilities also cannot perform well in practical subjects and the entire evaluation system had to change in a course of the year. Students pursuing bachelors and masters have been hit the worst, as in colleges there are students from various parts of the country. Students cannot always be available and face immense connectivity problem, from online class presentations to strenuous online exams it has become difficult for the students to be able to complete their academic commitments. Students who have changed their streams are clueless and do not know how to cope up with the syllabus and whom to ask, there have been no or minimal senior-junior interaction which possess even more challenge for the students. As masters or bachelors are the base of the professional world, in this if the students are not getting the proper exposure and experience the recruitment rate will fall eventually, currently only those organizations are hiring where work can be done in online mode for other practical exposure courses there are very less opportunities. For example, students pursuing masters or bachelors in social work are not getting the opportunity to work in the grass root levels and have lost many job opportunities as well, organizations working for these causes are now funding the pharmaceutical companies and vaccine drives leaving other aspects void for the moment. In these trying times, even the teachers are exhausted as they are not able to teach the way they used to and are facing challenges coping up with the digitized way of teaching, many school and college teachers are not aware of the technological tools in many online teaching platforms and they also face immense power cuts and connectivity issues, especially when they teach from their hometown(s). It has become challenging and tiring to evaluate the students and grade them in online performance as it is not easy to understand how much the student has actually been able to follow and whether the work is authentic or not. Female educators have to look after their families and teach their own children as well, responsibilities of family and school/college come together for these educators which makes it tougher for them to deliver their best as well. Thus, the online mode of education can be a temporary solution but continuing like this will weaken the base of the students and they cannot ever learn practical subjects or get field exposure that helps them in securing suitable jobs.

- Health Implications:
  Prolonged screen time and no or no or minimal exercise have exposed students to various health problems and most of the students are getting higher power glasses as well, entire work-from-home concept has caused devastating health effects on everyone. The youngsters are getting back pain and spinal cord pain which is not appreciated at such an age. Due to no physical activity school-going students and adolescents are getting hyperactive which is causing them to behave in certain abnormal ways, female homemakers have faced a lot of pressure once everything became online as they had to take care of everyone in the home from elderly parents to tender toddlers without any external help, the entire day's work caused numerous health issues as well like fatigue, body pain, and muscle pain, in such situation it is not possible for only one person to carry out all the household chores. People sitting in front of the screen all day long face problems like dizziness,

headache, vision problem, mild throbbing in a particular part of their head, and causing multiple pain areas. It is taking a toll on the people involved in office work which has increased due to the pandemic and the pressure to survive in this cut throat competition too is hounding the employees for performing well. From prolonged timings to early deadlines it has become very difficult to be able to keep track of health and other ailing issues; due to all these people have been neglecting their persisting health issues and have not been able to go to doctors for regular check-ups as well.

- Depleting Mental Health:
  Mental health now in the aspect of life is depleting and from toddlers to adults no one has been spared from the havoc. People all over the world are going through tremendous mental pressure and in this situation it has become impossible to maintain mental peace and carry on with everyday tasks. Situation is so bad that people have gone to the cremation center finding no vacant slot and stood alone with their loved ones' deceased bodies. Humanity has witnessed so many deaths and Covid-19 cases which has made everyone helpless and has forced them to live in negativity. From loss of lives to loss of jobs, from ill-maintained medical infrastructure to a mockery of education, the pandemic has made us through everything. It has become practically impossible to think that life will be normal ever again. Students cannot go out and study they have to rely on computers and laptops only, working professionals cannot assess the real ground level work which is affecting the people under them apart from IT sector professionals. It has all become just a virtual world where human touch is now scared of and fear of getting exposed to the virus is always there in the hearts and minds of everyone. Women homemakers are also not spared from this nightmare, in this situation they have to take care of their families and in most Indian families aged and children live which becomes a challenge for the female as she has to do all the household chores, house-help being unavailable and taking care of aged parents and school-going children. First wave made us all worry about the future but the second wave ripped us all apart. Few states in India were in such a situation where one in every three tests was positive and all the members of the family were affected by the virus. Seeing people dying in front of one's eye is not a thing to process easily, mental toll today is more than a war, we all are in the battlefield even being at home. People in rural India are now out of the communication circle which leaves them at the edge and they don't have the required medium of communication as well, it has made them vulnerable and without any income, as there still is inter-state travel restrictions they cannot come to big cities to resume their work and earn their daily wage, they cannot undergo state-mandated quarantine as the cost has to bore by the people traveling only. Pandemic sure is a black-swan event that has put a halt on everyday life and only the most privileged are unaffected and indifferent. Students, adolescents, young professionals, young mothers, and elderly people from every walk of life have been affected mentally and cannot live the way they had wished. Every aspect of life has been stopped in some or other way students and young professionals are committing suicides more than ever which is an alarming matter to look into, lack

of availability of resources and equipment have forced many workers to search for meager jobs which barely support their families. In this constant state of confusion and uneasiness life has become extremely challenging for small business owners, street hawkers, restaurant owners, hospitality sector, tourism, and the list never ends. Anxiety, helplessness, trauma, depression, unhappiness, and confusion has become synonymous with one's survival nowadays, people who have recovered from the disease face issues like extreme heart palpitation, indigestion, headache, and other issues to which doctors also have no proper solution, such a state of mind leads to anxiety and tension which further depletes one's mental health. Solution to a disease is a drug/medicine, vaccination boost the immune system in a body but does not guarantee relief from the disease, people are so much disturbed and disheartened that they fear to discuss their issues with anyone and feel they can handle everything on their own which technically is not possible. In these trying times, people should stand for each other in any way possible so that no one loses hope and keep on trying their best until the black phase is finally over.

3 Proposed Solution

The objective of this solution framework is to build an Artificial Intelligence-enabled robust digital framework that is capable of collating different sources of information for forecasting the next wave at a more granular level. This will help the nation to effectively combat the challenges and unprecedented threats posed by this invisible pandemic war. This pandemic has not only impacted the population through the virus incursion but also due to economic and mental collapse, the developing countries are suffering from unemployment and hunger crises. Our solution framework will also help in identifying the best suitable employment opportunities for the impacted individuals.

Today's state-of-the-art forecasting tools use machine learning, a type of AI that relies on historical data to make predictions for the future. But unlike recurring epidemics, which provide useful information in their wake—pandemics offer little historical data to learn from. By definition, a pandemic is the worldwide spread of a new disease. That means that, in the beginning, there is no availability of historical data which can be leveraged to build and train the model. To tackle these challenges, we propose an Autonomous and incremental artificial intelligence framework that will be capable of adapting to the new behaviors using multivariate point process modeling. Our underlying hypothesis behind this experiment is, that when the events occurring over a certain period of time are stochastically inhibited, the sequence of events will be playing a crucial role in deciding the likelihood of the future event. The point process modeling is highly flexible in accommodating demographic features along with mobility trends, which are connected using statistical models. This will help us in forecasting the trend at a more granular level. This will help the authorities to ***take*** proactive pivotal measures in advance to combat the fight against Covid-19, by scaling up the medical care facilities and optimizing the logistics effort. While uncertainties and changes are an inescapable part of this pandemic, our proposed framework will be playing the role of a helping aid in this catastrophic pandemic management from here to eternity.

Every year, more than 120 million workers across India migrate from rural areas to larger cities for employment opportunities. During the lockdown, these migrant workers who wanted to return home often struggled for food, shelter, transport, and employment opportunities. Although humanitarian aids and supporting offers from government agencies, non-profitable enterprises were offered from every corner of the country to meet their most-pressing needs, connecting the help providers with the people in need at scale seems to be the biggest challenge.

Our proposed framework is capable of enabling the unemployed workers community to locate the resources closest to them using machine learning algorithms streaming real-time information at scale. The algorithm attempts to find the optimal similarity matrix between the workers and the openings available through semi-supervised learning using KNN-based Label propagation and Label spreading approaches and adaptive PU learning. The algorithm further adds a layer of regularization to be more robust toward the noise.

4 Solution Architecture

The proposed architecture consists of five layers stacked over one another. Layer one will be responsible for accumulating all sources of information arriving from the governing authorities. The next layer contains AI-powered point process algorithm. The algorithm acts as the backbone of the framework. It helps us analyze every single information feed arriving into the system with utmost granularity. Layer three acts as the response layer, forecasting the next hit and assisting in proactive measures of precaution and safety. Layer four is an extensive wing of the framework specially designed for tackling sociological challenges faced by humanity during this pandemic. This pandemic has taken a big toll on daily wage earners, the unprivileged, and the homeless. The machine will be acting as a helping hand to those in need using optimized nearest neighbors propagation and utility spreading approaches at scale.

The concluding layer holds the responsibility of analyzing the driver along with monitoring and vigilance, which will be backpropagated to the layer two and layer four algorithmic framework for incremental learning and proactive refinements improving the data-centric decision driving capabilities.

The framework architecture of the proposed solution can be found in Fig. 1. Diverse sources of data will be considered for accumulating feeds such as feeds from the governing authorities regarding different demographical information, domestic and international border activities, social media extracts, hospital occupancy data, travel itinerary, extracts from GPS, Smart devices, WIFI, etc. These information feeds will be further processed with sparsity prediction/removal, outlier management, and different data wrangling processes using state-of-the-art exploratory data analysis and cleansing algorithms. Next, they will be stored in data centers in-house/cloud

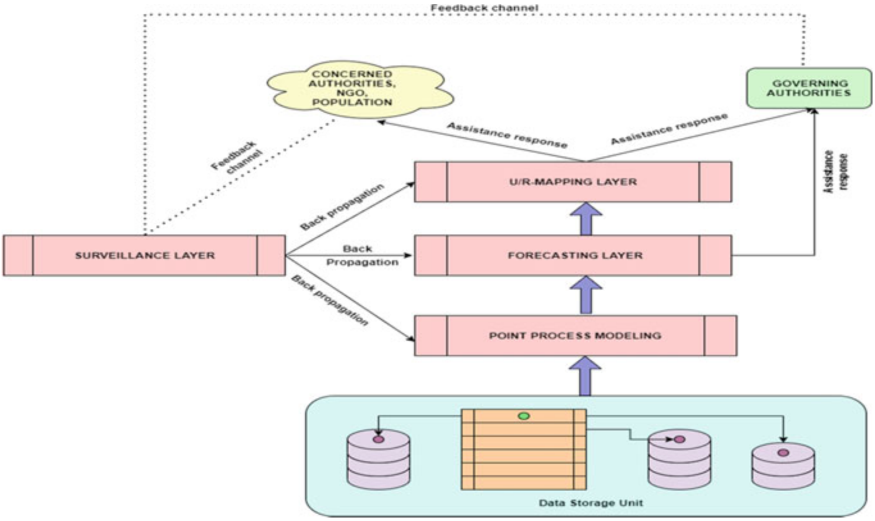

Fig. 1 Proposed Solution Architecture

storage. These storage units are capable of storing heterogeneous data, which can be employed to store any type of data objects.

4.1 Point Process Modeling: Hawkes Process

Hawkes process was first introduced by Alan Hawkes in 1971. It is a statistical modeling pipeline, build on self-exciting methods where the probability of an event's occurrence is dependent on the likelihood of another event's occurrence. This basically shows how on one particular event other events are dependent and follows certain patterns to give the most appropriate result, it is a process which is widely used to assess the occurrence of crimes due to various factors, health of people and susceptibility to diseases, market analysis and in various other fields. It helps to predict future actions and how to control certain disasters by taking appropriate measures beforehand.

Hawkes process is a statistical framework for building self-exciting modules, used for building the pipeline to predict the trend and then make users or authorities alert for the future course of action, it follows certain data and gives the most appropriate result, there are few Hawkes Models which are used to solve the complex situations and it helps in accurate prediction of the scenario when provided with required data.

Understanding Various Hawkes Processes in Prediction of Covid Hotspot Generation:

- Hawkes Model
  The intensity function of the Hawkes pipeline is defined as:-

$$\lambda(t) = \mu + \sum_{j:t_j<t} \psi\left(t - t_j\right) \tag{1}$$

  In the above equation, μ is the base intensity and $\psi(\cdot)$ is a pre-specified decaying function, i.e., exponential function and power-law function. Naturally, Eq. 1 means that each of the past events of contraction with Covid-19 virus has a positive contribution to the occurrence of the current event of generation of a new hotspot area, and this influence decreases with the passage of time.
- Intensity Function:
  The intensity function [1] is assumed that all the recurring events associated with a number of cases and other underlying events affect the probability of the formation of a new hotspot. This helps in leveraging the information of the historical and co-occurring events like incoming travelers, migrant labors, people who contracted with the virus that have been updated/reported. The intensity function is considered to be representative of the baseline intensity and the densities of different the historical or co-occurring events of contraction with the virus which might lead to another hotspot generation.

$$\lambda_{y,m}(t) = \mu_y + \sum_{t_\ell<t} \mathbb{I}(m_\ell = m)\alpha_{y_\ell,y}\kappa(t - t_\ell) \tag{2}$$

  Here the first term represents the constant base intensity of generating label y. The second term represents the influence of the events and the cases reported that happen prior to the time of interest. The influence from each event's occurrence decays over time and is modeled using an exponential decay term $\kappa(t - t_\ell) = \omega \exp(-\omega(t - t_\ell))$. The matrix α of dimensionality $|Y| \times |Y|$ represents the encoding of the level of dependency between pairs of areas belonging to the events related to getting contracted with the Covid-19 virus, e.g., questioning an area's trend (increasing/decreasing) may influence the chances of rejecting situation in the future generation of a hotspot area differently from how it would influence a new area where more people are traveling.
- Likelihood Function:
  The different set of parameters present in the intensity function gets optimistically decided by maximizing the likelihood of the events of contraction with the virus. The complete likelihood function is given by

$$\begin{gathered} L(t, y, m, W) = \\ \prod_{n=1}^{N} p(\boldsymbol{W}_n \mid y_n) \times \left[\prod_{n=1}^{N} \lambda_{y_n,m_n}(t_n)\right] \times p(E_T) \end{gathered} \tag{3}$$

Here the first term provides the likelihood of generating a spike in the number of cases given the area and is modeled as a multinomial distribution [2] conditioned on the area.

$$p(\boldsymbol{W}_n \mid y_n) = \prod_{v=1}^{V} \beta_{y_n v}^{W_{nv}} \tag{4}$$

Here V is the population sample size and β is the matrix of size |Y| × V specifying the point process model for each label. The second term provides the probability dynamics in the number of cases at times $[t_1, \ldots, t_n]$ and the third term provides likelihood in a way that the probability density curve flattens at [0, T] timestamp, with the exceptions incurred at times $[t_1, \ldots, t_n]$. The parameters of the model get optimized by the log-likelihood maximization.

$$\begin{aligned} l(t, y, m, W) = -\sum_{y=1}^{|Y|} \sum_{m=1}^{|D|} \int_0^T \lambda_{y,m}(s)ds + \\ \sum_{n=1}^{N} log\lambda_{y_n,m_n}(t_n) + \sum_{n=1}^{N} \sum_{v=1}^{V} W_{nv} log\beta_{y_n v} \end{aligned} \tag{5}$$

The term of integration in Eq. (5) gets computed for the function of intensity as the constant function and exponential decay function are integrable.

Here, point to remember is that β is not dependent on the dynamics part, and gets derived from solution post Laplacian smoothing application.

$$\beta_{yv} = \frac{\sum_{n=1}^{N} \mathbb{I}(y_n = y) W_{nv} + 1}{\sum_{n=1}^{N} \sum_{v=1}^{V} \mathbb{I}(y_n = y) W_{nv} + V}$$

For α and μ optimization method (HP Approx.), approximation is taken that the logarithm in Eq. (5) by calculating the log within the summation in Eq. (2).

$$\begin{aligned} \mu_y &= \frac{\sum_{n=1}^{N} \mathbb{I}(y_n = y)}{T|D|} \\ \alpha_{ij} &= \frac{\sum_{n=1}^{N} \sum_{l=1}^{n} \mathbb{I}(m_l = m_n)\mathbb{I}(y_l = i)\mathbb{I}(y_n = j)}{\sum_{k=1}^{N} \mathbb{I}(y_k = i) K(T - t_k)} \end{aligned}$$

In the above equation $K(T - t_k) = 1 - exp(-\omega(T - t_k))$ gets calculated by integrating $\kappa(t - t_k)$. In an alternate approach, it is found that parameters using joint gradient-based optimization over μ and α, using derivatives of log-likelihood $\frac{dl}{d\mu}$ and $\frac{dl}{d\alpha}$. In optimization, the process operates in the logspace of the parameters in order to ensure positivity, and employ L-BFGS approach to gradient search. Moreover, it initializes parameters with those found by the HP Approx. method. It is used to normalize the decay parameter ω, in our case to 0.1.

- Transformer Point Process Modeling:
  Transformer point process modeling [3] is an attention centric pipeline that is broadly applied to translation modeling applications and NLP.

  The transformer model can be interpreted with the help of a series of examples, assuming an event sequence, $\mathcal{S} = \{(t_j, k_j)\}_{j=1}^{P}$ comprises P events, and every event occurrence belongs to type $pj \in \{1,2,\ldots,P\}$, with aggregated occurrences of P types. Every pair (t_j, p_j) belongs to an p_j type of event, which take place at t_j time.

  The ground hypothesis of transformer-based point process modeling is to attain a self-attention module. It is quite different from RNNs, as the attention module does not follow the recurrent structures. Although the model handles the temporal dynamics of the inputs provided, using time distributed dynamics and batch processing. Hence, similar to the positional encoding methodology the usage of a temporal encoding method is recommended for obtaining better results.

  Figure 2 provides an explanation of how each event sequence S gets connected with the embedding layers, followed by the self-attention modules. The output will be embeddings generated of event S, with knowledge of prior events encoded inside as a knowledge layer.

  The underlying equation summarizes the story behind the process:

$$[\mathbf{z}(t_j)]_i = \begin{cases} cos\left(t_j/10000^{\frac{i-1}{M}}\right), \text{ if } i \text{ is odd,} \\ sin\left(t_j/10000^{\frac{i}{M}}\right), \quad \text{if } i \text{ is even.} \end{cases} \tag{6}$$

  The model utilizes trigonometric equations for building a temporal encoding layer at a granularity of every time stamp obtained. At every t_j, the module deterministically generates $\mathbf{z}(t_j) \in \mathbb{R}^M$, where M is the encoding dimension. It has

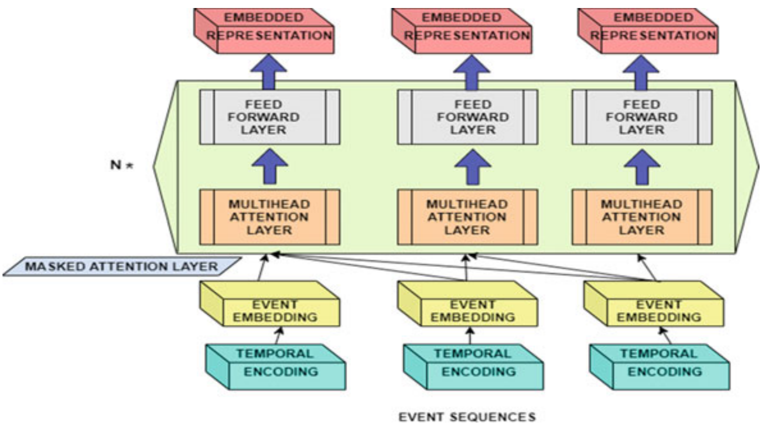

Fig. 2 Transformer Point Process Model

the feasibility of employing several encoding methods, like the position relativity representation module, where instead of predefining, two temporal encoding matrices are generated. The model has the feasibility of training the matrix of embedding $U \in \mathbb{R}^{M*R}$ for the type of events, where the r^{-}th column of U represents an embedding of dimension M for r type of event. Considering an event r_j, with the one-hot encoded (a vector space of r-dimensions with 1 s for the r_j-th point, and the rest of the indexes contains 0 s), embedding as Ur_j. For an event with the relative time stamp (t_j, r_j), the event collation Ur_j and the temporal embedding $z(t_j)$ resides in RM. The embedded representation of the event sequence $\mathcal{S} = \{(t_j, r_j)\}_{j=1}^{L}$ can be specified as follows:

$$X = (UY + Z)^T \tag{7}$$

In the equation [4] mentioned above $\mathbf{Y} = [\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_L] \in \mathbb{R}^{K\times L}$ is the collective event types, their encodings, $\mathbf{Z} = [\mathbf{z}(t_1), \mathbf{z}(t_2), \ldots, \mathbf{z}(t_L)] \in \mathbb{R}^{M\times L}$ and is the cumulative encodings of different event types. Here the key insight is $X \in \mathbb{R}^{L*M}$ and the entries of X belong to the embedded event in order of their relativity. Next X will be passed through the self-attention layer, and output S is formulated as-

$$\mathbf{S} = \mathrm{Softmax}\left(\frac{\mathrm{QK}^{\top}}{\sqrt{M_K}}\right)\mathbf{V}, \text{ where } \mathbf{Q} = \mathbf{X}\mathbf{W}^{Q}, \mathbf{K} = \mathbf{X}\mathbf{W}^{K}, \mathbf{V} = \mathbf{X}\mathbf{W}^{V} \tag{8}$$

In Eq. 9, {K},{Q}, and {V} represent the key, query, and value matrices generated from various transformations of X, and $\mathbf{W}^{Q}, \mathbf{W}^{K} \in \mathbb{R}^{M\times M_K}, \mathbf{W}^{V} \in \mathbb{R}^{M\times M_V}$ are the coefficients of relative importance of numerous linear transformations, performed in order of their relativity. The usage of multi-head self-attention is recommended for the reduction of input feature space. Various outputs from the attention layer $S_1, S_2, \ldots, S_H$ are generated using weights $\left\{\mathbf{W}_h^{Q}, \mathbf{W}_h^{K}, \mathbf{W}_h^{V}\right\}_{h=1}^{H}$. The final outcome can be generated as follows:

$$\mathrm{S} = [\mathrm{S}_1, \mathrm{S}_2, \mathrm{S}_3, \ldots.., \mathrm{S}_\mathrm{H}]\mathrm{W}^{O} \tag{9}$$

In Eq. 10, $\mathbf{W}^{O} \in \mathbb{R}^{HM_V\times M}$ is the accumulation matrix. The self-attention module identifies the events having their occurrence, at a certain distance from the present time space. Column j, in the attention layer weights Softmax $\left(\mathbf{QK}^{\top}/\sqrt{M_K}\right)$ denotes the dependency of event t_j on historical event occurrences, whereas Recurrent neural network models encode history information in relative sequence, i.e., t_j state has dependencies on t_{j-1}, which eventually has a dependency on the t_{j-2} state and so on.

Outputs of the attention layer is next propagated through a feed-forward neural network architecture, where it generates encrypted representations q(t), provided as input to the layer:

$$\mathbf{H} = \mathrm{ReLU}\left(\mathbf{S}\mathbf{W}_1^{\mathrm{FC}} + \mathbf{b}_1\right)\mathbf{W}_2^{\mathrm{FC}} + \mathbf{b}_2, \mathbf{q}\left(t_j\right) = \mathbf{H}(j, :) \tag{10}$$

In Eq. 11, $\mathbf{W}_1^{\mathrm{FC}} \in \mathbb{R}^{M\times M_H}$, $\mathbf{W}_2^{\mathrm{FC}} \in \mathbb{R}^{M_H\times M}$, $\mathbf{b}_1 \in \mathbb{R}^{M_H}$, and $\mathbf{b}_2 \in \mathbb{R}^{M}$ represents the hyper parameters constituting the neural network, and WFC 2 has similar set of columns. The resultant metric H $\in \mathbb{R}^{LXM}$ provides embedding's of the different event occurrences in the input space, where each entry belongs to an individual event. While the computation of the output from the attention layer S(j,:) (the j-th row of S), look forward positions have been masked. It will be helpful in avoiding the dependency on look forward events. Figure 2, illustrates the architecture of Transformer Hawkes Model.

- Continuous Time Conditioning Intensity:
  Dynamics of event occurrences is formulated as a continuous time conditioning intensity function. Equation 10 produces the embeddings for different time stamps, and the intensity associated also follows the discrete distribution. Considering $\lambda(t|H_t)$ as the model's conditional intensity metric, where $H_t = \{(t_j,k_j): t_j < t\}$ represents the historical events up to time t, it can be derived from the various intensity functions corresponding to specific event types. Every k $\in$ {1,2,…,K}, formulates $\lambda_k(t|H_t)$ to be the conditional probability density function, for event occurrences of type k. and can be defined as follows:

  $$\lambda(t \mid \mathcal{H}_t) = \sum_{k=1}^{K} \lambda_k(t \mid \mathcal{H}_t)$$

  Here the intensity function of a specific form can be represented as:

  $$\lambda_k\left(t \middle| \mathcal{H}_t\right) = f_k\left(\underbrace{\alpha_k \frac{t-t_j}{t_j}}_{\text{current}} + \underbrace{\mathbf{w}_k^{\top}\mathbf{h}\left(t_j\right)}_{\text{history}} + \underbrace{b_k}_{\text{base}}\right) \tag{11}$$

  Here, the time interval $t \in \left[t_j, t_{j+1}\right)$, and $f_k(x) = \beta_k \log(1+\exp(x/\beta_k))$ represents the softplus function with "softness" hyper parameter as β_k. The selection of the function can be explained with twofold justification: the softplus function provides the assurance of positive intensity; and the "softness" of the function ensures the stability of computation, avoiding the chances of dramatic fluctuations in the intensity.

  - The "current" influence [5] represents the degree of influence between multiple time points t_j and t_{j+1}, and α_k signifies the relative importance associated with them. If $t = t_j$, then a new observation arrives in, and its influence will be 0. If $t \to t_{j+1}$, the conditional intensity function loses its nature of continuity.
  - The "history" term comprises dual significance [6]: the W_k vector is responsible for transformations in the embedded states of the Transformer Point Process

Pipeline converts to a scalar entity, and the embedding *h(t)* encodes past events till time space t.

The "baseline" intensity signifies the likelihood of an event incident without taking into consideration the dynamics of historical event occurrences.

4.2 Label Propagation Algorithm

Label propagation was originally introduced by Xiaojin Zhu and Zoubin Ghahramani in the year 2002. This is a semi-supervised algorithm which will be helping the needy population to locate both the financial and medical resources nearest to them without any hassle. The algorithm behind the framework is as follows.

The framework forms a graph architecture [7] with the resource locating feeds as nodes. Next it puts a weighted edge between the nodes using time-centric distance calculation and edge detection algorithms. The framework utilizes dynamic time wrapping-based algorithms for achieving calculating the similarity between the journeys. We will be explaining the process with an example. Assuming we have a sample dataset (2D) that consists of only a few samples, as presented in Fig. 3. There exists binary classified feeds and one among them is un-labeled. The samples represent different nodes of the graph. The framework will be connecting each node with any other node based on dynamic time wrapping-based similarity calculation, and annotate the edges with the resultant distances (Fig. 4).

The algorithm will assign higher weights to the nodes lying closer to each other through the Gaussian transformation using the radial basis function [8].

$$w(x, x') = e^{-\left(\frac{x-x'}{\sigma}\right)^2} \tag{12}$$

Here x and x' are the resources located in the surrounding space. If the consumers and the resources are really close, i.e., $|x\text{-}x'|$ is close to 0, their assigned weightage [9]

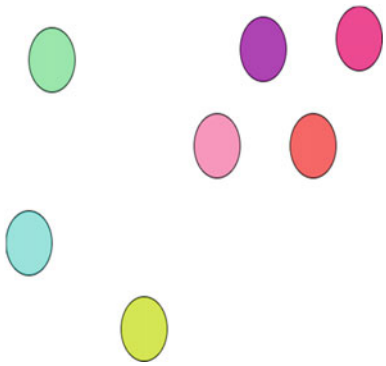

Fig. 3 Data points in n dimensional space

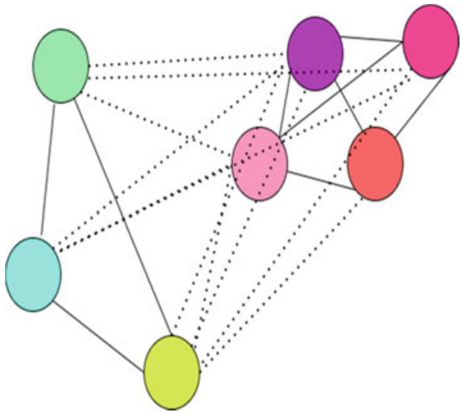

Fig. 4 Label Propagation based on similarity

will be approaching 1 and in case they are further apart from each other, the weight will tend to approach 0.

σ is a hyperparameter to be optimized over time based on model performances and loss reduction in allocation efficiency.

Next for locating the nearest resource existing from an individual, the framework will start a random walk originating from its locus. One step of the walk refers to the movement from a particular node to another one except the source node itself. Edges associated with a higher weightage will be prioritized for optimal efficiency. It applies Markov chain modeling along with K nearest neighbors-based algorithms for setting up the priorities between the crosswalk selections. Markov chains are a stochastic modeling method that explains the occurrence of a series of possible events where the likelihood of an individual event depends only on the state attained from the previous event maintaining the sequence history for data-centric decision building. Finally, the label of each of the points will be assigned to the nearest neighboring resources.

4.3 Framework Modeling

A point process model is a collection of randomly located points located on a line or Euclidean space. The modeling algorithm is a stochastic process underlying the occurrences of events in a particular time/space. The occurrences of every single event including different pandemics like Covid-19 are very likely to follow a particular trend, and our intention will be to capture and analyze such patterns at the utmost granular level, so that those learnings can be utilized in the future to prevent or fight against such pandemics. For example likelihood of a person getting a severe heart attack increases when the person has high blood pressure and belongs to the elderly age groups than in his middle ages with stable blood pressure. The panic created due to instability of sales and revenue generation of a corporate firm in one country can cause similar events in other countries, which will further lead to a disaster in the public valuation of the company. Taking an example of the Amazon forest that suffered from wild file this year, created awareness in the entire world which will eventually reduce the likelihood of another forest, facing wildfire going forward. These examples helped us understand the fact that the likelihood of the occurrence of an event can be increased or decreased following the patterns in the sequence of prior events happening around the world.

If the events happened to increase the likelihood of the occurrence of the future event, those events are known as stochastically [10] excited or self-excited events. The example of a person having a heart attack falls under this category, whereas if the likelihood of a similar event is decreased like the example of the impact of company valuation, then it is known as stochastically inhibited or self-regulating events (Fig. 5).

The proposed solution intends to capture and analyze these patterns present in the surveillance feeds, collected from the governing authorities. Once the patterns are identified they can be clustered based on comprehensive dynamic time wrapping distance measurement. The proposed framework [11] intends to capture all possible sources of information, granulated at minimum source level to capture live information feeds capturing demographical, hospitality, and mobility metrics. The model will identify the chaining pattern existing between the events, proactively forecast the future trend before time and assist governing authorities with alarming alerts, in case of a requirement for additional care. The framework is capable of sending out alarming alerts via different mediums. This pipeline of processes will eventually help in resource optimization and strategic execution. The forecasting layer also outputs regular monitoring status checks along with forecasted status predictions for the next 90 days, enabling greater visibility and control over the pandemic (Fig. 6).

During the catastrophic pandemic, there has been a severe crisis in employment opportunities for the working class. Although humanitarian aids and supporting offers

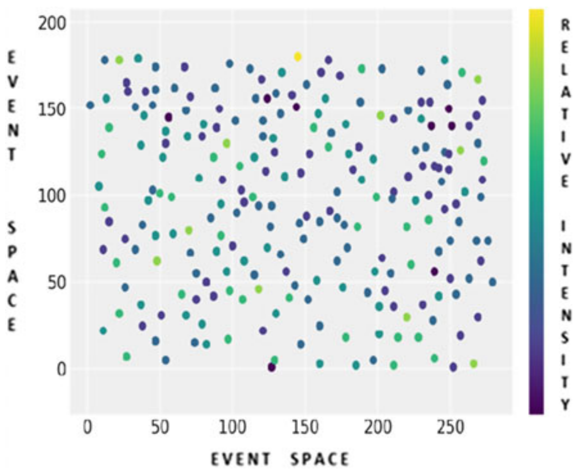

Fig. 5 Relative Intensities among the events

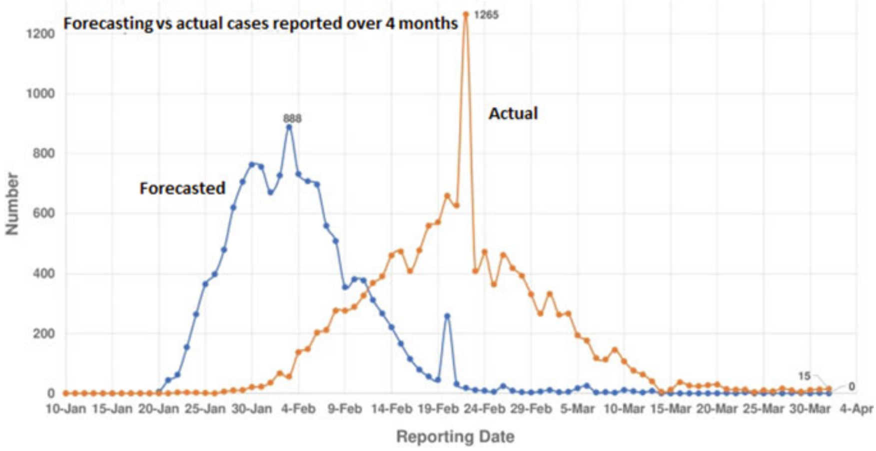

Fig. 6 Actual vs Forecasting

from government agencies, non-profitable enterprises were offered from every corner of the country to meet their most-pressing needs, connecting the help providers with the people in need at scale seems to be the biggest challenge. The proposed solution (U/R Mapping layer) [12] will enable the population to locate the resources/utilities closest to them at their fingertip using a semi-supervised learning algorithm. The algorithm is inspired by a technique from experimental psychology called spreading activation networks. The algorithm employs the points, present in the dataset to be connected in a graph-based network on their relative distances in the input space. The weight matrix of the graph is then normalized symmetrically, much like spectral clustering methods. Information is passed through the graph, which is adapted to capture the structure in the input space. Finally, the label of each un-labeled point is set to be the class from which it has received the most information during the iteration process (Fig. 7).

The surveillance layer is responsible for tracking the performance of model prediction and tries to learn from the mistakes, done in the past. The base idea behind this is to make the pipeline learn from the environment by interacting with it and receiving rewards for performing correct actions. Interaction with the environment stands for gathering the event occurrence pattern over a period of time and analyzing the hidden trends present in the data.

$$\text{Reward function} = \frac{(0.75 * \text{ previous batch count}) + (0.85 * \text{ current batch count})}{(0.75 * \text{ previous batch size}) + (0.85 * \text{ current batch size})} \tag{13}$$

The algorithm will refine with time using incremental learning based exponential smoothing based reward function mentioned in Eq. 13. The key insights derived will be backpropagated to the modeling algorithms to penalize the outdated patterns and

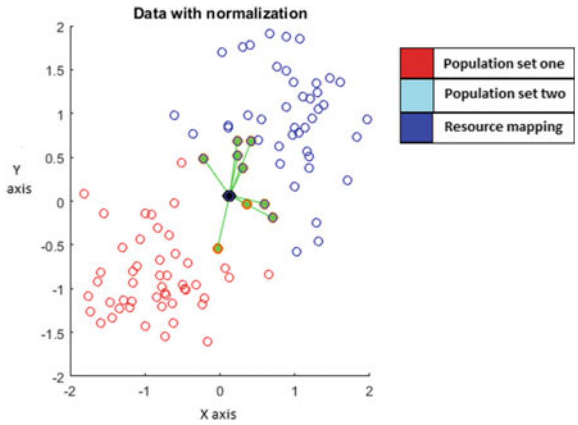

Fig. 7 Label Propagation

promote the recent trends to penalize the outdated patterns and promote the recent trends.

5 Conclusion

In this book chapter, we have introduced an Artificial Intelligence based framework to tackle pandemics like Covid-19. We have applied point process modeling using the Hawkes process [13] for the next Covid hotspot prediction, which will help the governing authorities to receive prior alerts and be prepared for the upcoming fights against these deadly pandemics and save humanity. This pandemic has ripped the world apart in every possible way from academic to financial, physical to physiological. The resource/utility mapping layer of the framework intends to solve the problem of employability and nearest food and medical resource allocation using dynamic time wrapping Label propagation methods. This pandemic has triggered one of the worst job crises in the entire world and it is being estimated that it will increase the poverty rate along with widening inequalities in the years to come. The governing authorities need to brainstorm collectively to take every possible action to stop these crises from turning into a serious social crisis. Our framework attempts to contribute to the same cause. It has been made capable of learning incrementally over time to adapt to the new environment and update itself for providing the best possible assistance toward re-establishing the normal.

References

1. Bacry, E., Mastromatteo, I., Muzy, J.-F.: Hawkes processes in finance. Market Microstruct. Liq. **1**, 1550005 (2015)

2. Bengio, Y., Simard, P., Frasconi, P.: Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Netw. **5**, 157–166 (1994)
3. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding (2018). arXiv preprint. arXiv:1810.04805
4. Hawkes, A.G.: Spectra of some self-exciting and mutually exciting point processes. Biometrika **58**, 83–90 (1971)
5. Li, S., Xiao, S., Zhu, S., Du, N., Xie, Y. and Song, L.: Learning temporal point processes via reinforcement learning. In: Advances in Neural Information Processing Systems (2018)
6. Linderman, S., Adams, R.: Discovering latent network structure in point process data. In: International Conference on Machine Learning (2014)
7. Mei, H., Eisner, J.M.: The neural hawkes process: A neurally self-modulating multivariate point process. In: Advances in Neural Information Processing Systems (2017)
8. Xiao, S., Yan, J., Yang, X., Zha, H., Chu, S. M.: Modeling the intensity function of point process via recurrent neural networks. In: Thirty-First AAAI Conference on Artificial Intelligence (2017)
9. Yang, S.-H., Long, B., Smola, A., Sadagopan, N., Zheng, Z., Zha, H.: Like like alike: joint friendship and interest propagation in social networks. In: Proceedings of the 20th International Conference on World Wide Web (2011)
10. Yin, W., Kann, K., Yu, M., Schutze, H.: Comparative study of cnn and rnn for natural language processing (2017). arXiv preprint. arXiv:1702.01923
11. Zhang, Q., Lipani, A., Kirnap, O., Yilmaz, E.: Self-attentive hawkes processes (2019). arXiv preprint. arXiv:1907.07561
12. Zhao, Q., Erdogdu, M. A., He, H. Y., Rajaraman, A., Leskovec, J.: Seismic: A selfexciting point process model for predicting tweet popularity. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM (2015)
13. Zhou, K., Zha, H., Song, L.: Learning social infectivity in sparse low-rank networks using multi-dimensional hawkes processes. In: Artificial Intelligence and Statistics (2013)