International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

IJIEMR Transactions, online available on 26th Apr 2022. Link
:http://www.ijiemr.org/downloads.php?vol=Volume-11&issue= Spl Issue 05

DOI: 10.48047/IJIEMR/V11/SPL ISSUE 05/08

Title **A Model for Hurricane Path Prediction**

Volume 11, SPL ISSUE 05, Pages: 44-55

Paper Authors
Mr. Meduri Sri Sai Krishna, Ms. Chelluri Pavani Suma, Ms. Kota Lokesh, Mr. Marri Durga Prasada Rao, Mr. Badavada Naveen Raju

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

A Model for Hurricane Path Prediction

Mr. Meduri Sri Sai Krishna[1], Ms. Chelluri Pavani Suma[2], Ms. Kota Lokesh[3], Mr. Marri Durga Prasada Rao[4], Mr. Badavada Naveen Raju[5]

[1]Assistant Professor, Dept. of CSE, [2]18ME1A0518, [3]18ME1A0542, [4] 18ME1A0557, [5] 18ME1A0512, Ramachandra College of Engineering, Eluru, A.P., India

msrisaikrishna21@gmail.com, pavanichelluri31@gmail.com,kotalokeswari123@gmail.com, mdpr666@gmail.com,naveenraju7799@gmail.com

Abstract - Property loss, personal danger, and civic upheaval have cost billions of dollars as a result of hurricanes. Hurricanes' devastating effects have prompted substantial research in hurricane forecasting, including discipline of storm trajectory forecast. The haversine distances are used to measure realistic hurricane trajectories. Traditionally, meteorological analysis was utilized to forecast a hurricane's future path. We provide an alternative strategy in this work. The haversine distances are used to measure realistic hurricane trajectories. We only used data from 1950 to 2000 to train our prediction model. Furthermore, we weight the training data to errand recent hurricane tendencies while decisive a trajectory forecast. Our model produced a prediction - correctness proportion of up to 10.0 percent developed than the present state-of-the-art, which is 75.0 percent. The haversine distances are used to measure realistic hurricane trajectories.

Keywords :- Big data; hurricane; trajectory; frequent sequences; frequent patterns; prediction; haversine distance

I. Introduction and Related Works

Big data are omnipresent nowadays. Massive amounts of important data with varying degrees of authenticity can be generated and collected from various data sources, counting the Internet of Things (IoT). As a result, big data management, IoT infrastructure, and data science for big data analysis and mining are in high request [1]. Meteorological data is an example of large data (e.g., hurricane trajectories). The haversine distances are used to measure realistic hurricane trajectories. All trajectories in the unique database are retained by directly translating the data (rather than dispensation it using AprioriAll). This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure that the database utilised in the forecast model always include storms and their organize arrangements, eliminating the essential for repetitive representing in following AprioriAll implementations. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

The haversine distances are used to measure realistic hurricane trajectories. Hurricanes are a kind of tropical storm that only occurs in the Atlantic basin. The Saffir-Simpson Hurricane Wind Scale1 divides hurricanes into five categories founded on wind speed and projected injury. A single hurricane in the United States has caused damages of up to US$157 billion [2]. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. The subfields of hurricane prediction analysis are I hurricane strength forecast [4, 5] and (ii) hurricane track prediction (which has been investigated in meteorological and geographic information system (GIS) [6, 7]. For example, approximately previous studies (e.g., [8]) offered clustering-based algorithms for predicting typhoon track, but their relevance to hurricanes is

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

unknown. Here necessitates the use of data mining to forecast storm paths.

Data mining is the procedure of extracting non-trivial information from data that is implicit, previously unknown, and possibly useful. We are primarily interested in sequential mining, which tries to find interesting association rules and chronologically-ordered common patterns that can be used to make forecasts [9] on upcoming data admissions, among other data mining activities.
Sequential mining techniques, for example, can be rummage-sale to produce training data aimed at a storm track forecast model (e.g., AprioriAll [10]).

By immediately translating the data, all trajectories in the original database are preserved (rather than processing it using AprioriAll). This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure that the database utilised in the prediction model always include storms and their coordinate sequences, eliminating the need for repetitive mapping in subsequent AprioriAll executions. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure

Dong et al. [11] suggested a hurricane trajectory forecast model based on data mining (HTPDM) that mines sequential data using a modified variation of AprioriAll. HTPDM uses AprioriAll to develop interesting association rules using historical storm trajectories from 1900 to 2000, which are then compared to test data made up of hurricane trajectories after 2001 to 2008. A particular testing trajectory is divided into two sections by HTPDM: I the initial trajectory Ti and (ii) the terminal trajectory Tt. Based on a predefined fitness function, Ti is compared to all connotation rule pasts until a suitably matching rule R is mines sequential data using a modified variation of AprioriAll. HTPDM uses AprioriAll to develop interesting association rules using historical storm trajectories from 1900 to 2000 identified. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. To evaluate if the prediction was right, the result of R is compared to Tt using pattern-matching algorithms. HTPDM's results revealed that the model achieved a best-case accuracy ratio of 65.0 percent. This signifies that 65.0 percent of all testing trajectories for which a matching rule was successfully generated were correct forecasts. When HTPDM was used to account for testing trajectories that could not be matched with any rule, the correctness ratio dropped to a worst-case value of 57.5 percent As a result, a reasonable inquiry is: Can we obtain great precision? As a result, we introduce the WARD-HTP (Weighted-Asset, Realistic-Distance Hurricane Trajectory Predictor), a new model that applies several tactics to improve the correctness ratio. WARD-HTP recovers the realism of the hurricane trajectory forecast model by reducing the impact of historical factors and favouring the impact of contemporary trends. As an example, the tactics used in our WARD- HTP end up with a correctness ratio of 75.0 percent, which is a 10% improvement over HTPDM. All trajectories in the original database are retained by directly translating the data (rather than processing it using AprioriAll). This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure that the database utilised in the prediction model always include storms and their coordinate sequences, eliminating the need for repetitive mapping in subsequent AprioriAll executions. The rest of the paper is structured as follows. Our WARD-HTP is described in the following two sections. Sections IV and V provide the evaluation and findings. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure.

II.Our weighted-Asset, Realistic-Distance hurricane Trajectory predictor(WARD-HTP)

HURDAT2, an updated The Atlantic hurricane database 2 is used in our The Atlantic Oceanographic & Meteorological Laboratory, which is part of the US National Oceanic & Atmospheric Administration, maintains a weighted-asset, realistic-distance hurricane trajectory predictor (WARD-HTP). (hurricane, trajectory)-pairs were created using the data in HURDAT2. The HURDAT2 hurricane paths in the Atlantic cover a huge area, with some spanning over 10,000 kilometres amid the first and last reported sites As a result, the coordinates that are recorded are limited to those that are recorded. between 0 and 50 degrees north latitude and 20 and 100 degrees west longitude. The south Atlantic basin is roughly the size of this degree range. Hurricanes are tracked as a series of (latitude, longitude)-points that do not perfectly match traditional (x, y)-coordinates since latitude and longitude coordinates are points placed upon a sphere. The requirement to change from Euclidean to Polar coordinates system has created challenges such as reliable distance estimation. The method that creates trajectory prediction possibilities in together the existing HTPDM model and our future WARD-HTP model is founded on AprioriAll, which generates association instructions of the form "Hurricanes located at P may later be positioned at Q." This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure.

All trajectories in the innovative database are retained by straight translating the data (rather than dispensation it using AprioriAll). This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure that the database utilised in the forecast model always include storms and their coordinate sequences, eliminating the need for repetitive charting in subsequent AprioriAll executions. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure.

As a preprocessing step before launching the hurricane trajectory model, the training data from HURDAT2 remained transformed hooked on a (hurricaneID, coordinate Sequence) database. After the initial mapping stage, the trajectory database is arranged as shown in Figure 1. The database just requires two dimensions (the hurricane ID and the sequence of latitude-longitude coordinatesHURDAT2 assigns an ID to each storm that corresponds to its location (the Atlantic), the index well-ordered by arrival in the hurricane day, and the year itself. Hurricane Katrina in 2005, for example, corresponds to AL122005. All trajectories in the original database are retained by directly translating the data (rather than processing it using AprioriAll). This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure that the database utilised in the forecast model always include storms and their coordinate sequences, eliminating the need for repetitive mapping in following AprioriAll implementations. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

HurricaneID	**LatLongSequence**
AL081976	∟(26.0:-84.0), (25.3:-83.3), (24.7:-82.7)∟
AL091976	∟(31.7:-68.2), (33.4:-67.5), (35.2:-66.4)∟
AL101976	∟(14.0:-48.0), (14.0:-49.3), (14.0:-50.6)∟
AL111976	∟(12.5:-37.5), (13.0:-39.0), (13.5:-40.5)∟

Figure 1. A sample of the hurricane-coordinates database that has been mapped.

The support (i.e., occurrence count) of frequent hurricane coordinate sequences is more than or equal to the minimal support threshold (minsup). AprioriAll's sequence creation stage, in general, creates regular Starting with single-item sequences, sort patterns of increasing length (corresponding to frequent single-point sequences). This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

The (hurricaneID, coordinate Sequence) database is scanned for recurrent single-point trajectories, which remain then translated hooked on possible two-point trajectories using an internal join. To identify which candidates satisfy minsup, the candidate two-point trajectories are compared to the database. The candidate (k+1)-point trajectories are generated by successively an inner join against the recurrent k-point trajectories. If any of the candidate sequences' subsequences are rare, they must be pruned after the sequence generation reaches level 3. When Apriori All is run entirely, it produces a gathering of all common storm trajectory patterns that satisfy minsup. Furthermore, Apriori All is employed to maintain the order of recorded hurricane coordinate sequences. This is an unintended consequence of the inner-join step. Consider the following two hurricane paths: T1 and T2. T1's L–1 terminal coordinates must match T2's L–1 starting coordinates if and only if L is the current level of AprioriAll being executed. we employ both T1 and T2 to build a new trajectory T3. AprioriAll's inner-join step simulates the required storm trajectory-matching in order to find common hurricane trajectory patterns. The second stage of AprioriAll is completed once all feasible frequent hurricane trajectory patterns have been determined. The rule-generation stage is the next critical step, and it is in charge of finding and documenting intriguing association rules of the kind "Hurricanes located at P may subsequently be found at Q." Candidates for storm trajectory pattern matching are the rules generated by Apriori All. All association rule antecedents are matched to hurricane paths. The hurricane terminal track is then predicted using the result of the closest match rule. The prediction is regarded correct if the association rule result matches the actual hurricane terminal trajectory. WARD-HTP forecasts future storm pathways using historically common hurricane trajectories. The raw recorded hurricane coordinates are represented in degrees to discretize (latitude, longitude)-points illustrated in Figure 1. The raw data's too fine-grained nature precludes it from producing meaningful frequent coordinate sequences. As a result, efforts must be made to discretize the coordinate system, allowing for the production of more coarse-grained frequent patterns. WARD-HTP discretizes recorded hurricane coordinates into roughly square blocks of a specified size, stated in degrees. The new trajectory (Block 2:Block -7), (Block 2:Block -7), (Block 2:Block -7), (Block 2:Block -7), (Block 2:Block -7), (Block 2:Block -7), (Block 2:Block -7), (Block 2:Block -7), (Block 2:Block -7), (Block 2:Block -7), In this section, duplicate coordinates are deleted. Because the degree-size of the discretization block is an integer, all coordinates are transferred to their lowest divisor in relation to the discretization size. It is now possible to map coordinates to their corresponding area block in a coarse-grained manner. By immediately translating the data, all trajectories in the original database are preserved (rather than processing it using AprioriAll) This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure . This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure that the database utilised in the prediction model always include storms and their coordinate sequences, eliminating the need for repetitive mapping in subsequent AprioriAll executions. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

The mines sequential data using a modified variation of AprioriAll. HTPDM uses AprioriAll to develop interesting association rules using historical storm trajectories from 1900 to 2000, which are then compared to test data made up of hurricane trajectories from 2001 to 2008.. A particular testing trajectory is divided into two sections by HTPDM: I the initial trajectory Ti and This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. (ii) the terminal trajectory Tt. Based on a predefined fitness function, Ti is compared to all association rule antecedents until a suitably matching rule R is identified. To evaluate if the prediction was right,

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

the result of R is compared to Tt using pattern-matching algorithms. HTPDM's results revealed that the model achieved a best-case accuracy ratio of 65.0 percent. This signifies that 65.0 percent of all testing trajectories for which a matching rule was successfully generated were correct forecasts. When HTPDM was used to account for testing trajectories that could not be matched with any rule, the correctness ratio dropped to a worst-case value of 57.5 percent As a result, a reasonable inquiry is: Can we obtain great precision? As a result, we introduce the WARD-HTP (Weighted-Asset, Realistic-Distance Hurricane Trajectory Predictor), a new model that applies several tactics to improve the correctness ratio. WARD-HTP improves the realism of the hurricane trajectory prediction model by reducing the impact of historical factors and favouring the impact of contemporary trends. As an example, the tactics used in our WARD- HTP end up with a correctness ratio of 75.0 percent, which is a 10% improvement over HTPDM. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

Consider Figure 2, which was made using Matplotlib [12], a Python 2D plotting/visualization tool, with a sample of HURDAT2's recorded hurricane tracks. Using region discretization, continuous storm coordinate sequences are broken down into discrete region sequences of roughly uniform size and shape. WARD-HTP employs a discretization block size of 5, which roughly corresponds to a 500 km × 500 km real-world block dimension. Because the mapping between coordinate degrees and kilometres can be erroneous, and longitudinal points are dependent on location [13], the block dimension size is approximate.

The estimated size of 500 kilometres was proved to be an average storm total diameter 3, allowing for coarse-grain hurricane trajectory descriptions without losing precision. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

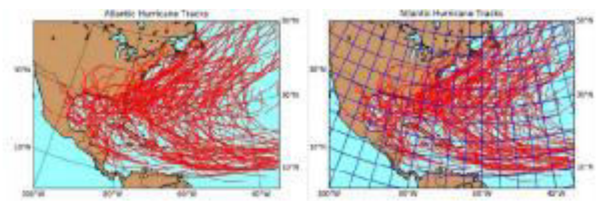

Figure 2. Hurricane trajectory paths (a) without and (b) with regionalization.

Region discretization, which scales all hurricane coordinate trajectories downward and allows the regionalized coordinates to be utilised as a simpler coordinate system in later sections of the prediction model, is required for hurricane trajectory prediction using data mining. Coordinate sequences that may become unusual in the future are included. Apriori All of the sequential mapping is carried out as a precaution. Property destruction, personal peril, and, in certain situations, severe societal discontent are all examples of damages.
[3]—prompt substantial research in the realm of storm forecasting. Then, while evaluating the best-fit rule for hurricane prediction, pattern matching is applied. The hurricane prediction model developed in this paper receives a new hurricane trajectory, which is divided into two sub-trajectories: I the beginning trajectory Ti and (ii) the terminal trajectory Tt. In this case, Ti is compared to the antecedent of every intriguing association rule that meets the match's minimum confidence threshold (minconf).The best-fit rule for an incoming trajectory that requires prediction is determined by two dimensions. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure . These measurements correspond to length and rule confidence. The longest corresponding

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

subsequence of decentralized points within the law antecedent within Ti is the matching length of the early trajectory Ti to a rule antecedent. Both Ti and the rule antecedent sequences have been regionalized using the methods described above. The best-fit instruction match for the incoming starting trajectory Ti is stated using a fitness function as follows:

$$\text{match} = \max(\text{conf}r \text{ x } (1 - \text{e}^{-\text{matchLen}(t,r)})), \ r \square AR \quad (1)$$

MatchLen(t, r) expresses the matching length between Ti and the antecedent sequence of r. The match formula is designed to be maximised in order to find the rule that best suits the impending hurricane path. After determining the rule of best fit, the result is compared to the actual hurricane terminal course to see if the prediction is right.Hurricane trajectory pattern matching is also used to assess the accuracy of predictions. The projected terminal sequence for the arriving hurricane trajectory is the result of the best-fit rule generated by the rule-fit matching stage. To measure the accuracy of the prediction, the best-fit result is compared to the actual terminal trajectory TtPrediction accuracy is a binary attribute in WARD-HTP: a prediction is either correct or incorrect. The forecast is valid if the matched-rule consequence also matches the impending storm terminal trajectory, and the model's correctness ratio is enhanced appropriately. In order to construct an real system for forecasting future storm paths, WARD-HTP uses modified AprioriAll, Downscales region discretization and pattern matching in terms of association rule antecedents and consequences. The model yields a best-case correctness ratio of 65.0 percent and a worst-case ratio of 57.5 percent based on Atlantic data.. storm trajectory data from 1900 to 2008. Trajectories from 1900 to 2000 were utilised as training data in the original model, with test data coming from 2001 to 2008. The model is constructed and tested at a high level by following the stages below and employing all of the implementation options discussed: From AprioriAll, generate interesting trajectory rules. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

1. Determine the incoming trajectory's matching rule. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.
2. Determine whether the matching rule's subsequent points are suitably close.

The capacity of an inward test trajectory to be coordinated with a trajectory connotation rule determines the difference between the model's best-case and worst-case hit rates. To find the best-fit match, the first sub-trajectory of an incoming trajectory is likened to all intriguing rules. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure.

III. Further Enhancement To OURWARD-HTP

A. *Haversine DistanceFormula*

We first offer a truthful distance calculation for storm trajectory comparison to further develop our prediction WARD-HTP. Remember that the Euclidean distance formula is calculated in two dimensions in HTPDM [11], and units are degrees of latitude and longitude., which correspond to latitude and longitude values, respectively.The Euclidean distance formula was created to calculate the unit distance between two N-dimensional planar locations. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

We utilise the haversine distance formula instead of the Euclidean method to calculate the On a vast sphere like the Earth's surface, the distance in kilometres between two latitude-longitude points. The haversine distance formula should be precise enough when applied to a specific location like the Atlantic Basin, despite the fact that the Earth's surface isn't a perfect sphere [13]. The two equations, denoted *D*Eucl and *D*havr are defined asfollows:

- □ *D*Eucl(*p, q*) = sqrt((□lat)^2 +(¯long)^2)
- □ Dhavr(*p,q*)=2*R*(atan2(sqrt(¯),s qrt(1–□)) where
- □ □=sin^2(0.5¯lat)+cos(*p*lat)+cos(*q*lat)+sin^2(0.5□long),
- □ □latand□long,respectively,denotethedifferencesin latitudeandlongitudebetweenpoints*p*and*q*.

It's worth noting that The Euclidean distance formula expresses distance in units on a plane with degrees of latitude and longitude, while the haversine formula expresses distance in portions of a single-unit sphere with just radians of latitude and longitude. As a result, multiply the value of the haversine formula by R, which equals the Earth's radius in kilometres, to represent distance in kilometres. (= 6,371 km [14]).In the discretized coordinate system, the Euclidean distance translates to a difference in square regions roughly equal to the stated region dimension. Assume the discretized coordinate system's region dimension is 5. A Euclidean distance of 1 between two 5-degree square regions R1 and R2 signifies that R1 and R2 are about 1 region (i.e., 5 degrees = 500 km) away after discretizing the coordinate system into square regions. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

B. In HTPDM [11], the Euclidean distance threshold is 1. When the haversine formula is used, the maximum distance varies between 300 and 450 kilometres, which is somewhat less than one average hurricane diameter. Because the average hurricane diameter4 is between 333 and 670 kilometres, two places are regarded inside the distance threshold if they are separated by less than one average hurricane diameter. The haversine formula's maximum distance ranges are substantially smaller than those employed in the Euclidean formula. As a result, when compared to findings derived using the Euclidean distance formula, the haversine formula is projected to produce finer-grained storm trajectory prediction predictions. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure.

C. Weighted TrainingData

We also use weighted training data in AprioriAll to improve trajectory prediction accuracy in WARD-HTP. AprioriAll favours recently formed trajectory patterns by enhancing their support by applying weights to recorded trajectories in a chronological order While the weighting of training data causes some trajectories to have a support greater than 1, this effect is minimised by counting previously recorded trajectories with a support less than 1. Increasing the support values for recent trajectories rather than decreasing the support values for early trajectories yielded the most promising results. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

Hurricane paths are weighed on a linearly rising scale in WARD-HTP. As a result, Before T, records were underweighted, whereas after T, records were overweighted.

Depending on the weighing method used, the total weight of all recordings does not have to equal 1.

Our WARD-HTP modifies AprioriAll further in

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

order to weigh the recorded hurricane records.
W specifies the total weighing scale to be used in the WARD-HTP modification. The last recorded hurricane track in the hurricane database HDB, for example, will have a support of 3 if W=3.

A 3/|HDB| assistance will also be given to the first trajectory.. For each entry in HDB the support modifier is increased by 3/|HDB|.The WARD-HTP ensures that newly created storm trajectory records are given more weight. Iterating through the HDB increases the current weight w by multiplying the weighting scale W by the HDB's size.As a result, the final record in the HDB will have a weighted support equal to W. Recent trajectories will require less repetition to develop association rules, whereas older data will require more repetition to build frequent sequences. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. The use of chronologically weighted data will result in association rules that favour current hurricane trajectory trends. As a result, the addition of weighted support to AprioriAll is projected to improve the accuracy of storm trajectory prediction for both current and future trajectories.

Figure 3. Non-matchable hurricane trajectory paths.

Year Interval for TrainingData

Third, we use recent data to train the model rather than historical data. While AprioriAll's weighted support helps to mitigate the impact of historical factors, the primitive nature of early storm recording technologies makes hurricane trajectories recorded before 1950 far less reliable than those recorded after [15]. The present storm trajectory prediction model is based solely on test data from trajectories recorded between 1950 and 2000. The goal of choosing this precise year interval was to try to reduce historical effects in the production of recurrent trajectory patterns (and subsequent interesting trajectory association rules).

All trajectories in the original database are retained by directly interpreting the data (rather than dispensation it using AprioriAll). This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure that the file utilised in the forecast model continuously include storms and their organize orders, eliminating the need for repetitive mapping in succeeding AprioriAll implementations.

Tweaking the training data year intermission, when combined with the haversine distance formula and weighted AprioriAll methods, is projected to result in a further rise in the model accuracy ratio. Furthermore, because historical influences will be removed, association rules created from training data from 1950 and onwards should be more reliable.

IV Evaluation

zWARD-HTP was fully implemented and tested against many input parameters to find the best configuration. We also experimented with different distance formulas, weighted training data methodologies (e.g., no weighted data), and training data year intervals. A Python implementation was used to test an 8-core Intel i7-4720 processor running at 3.35 GHz with 8 GB of RAM. We used HURDAT2 data for both training and testing.We do not attempt to extrapolate points for storm paths for which no rules were successfully met, which is a fundamental difference between the existing HTPDM [11] and our WARD-HTP. Our findings reveal that WARD- HTP accurately represents the accuracy ratio when considering both all test trajectories and only those with which a rule was immediately successfully matched. Figure 3 depicts the characteristics of trajectories that were unable to match with any meaningful criteria, using training data from 1950 to 2000 and test data from 2001 to 2015. A portion of the training data is represented by red trajectories, whereas blue trajectories are those that did not meet any association criteria. The non-matching trajectories were clearly distinguishable from the training data trajectories. As a result, no amount of point extrapolation can produce a valid terminal trajectory forecast. Some of the non-matching trajectories appear to be aligned with the training trajectories, but they are travelling in the other direction. As a result, any locations that matched such trajectories would not yield a trustworthy result.

When comparing WARD-HTP to HTPDM, the haversine distance, weighted training data, and training data year period of 1950-2000 all lead to a total increase in best-case accuracy ratio of 10.0. percent, for a total of 75.0 percent. When set in the same way that produced the best-case result of 75.0 percent, the worst-case correctness ratio (ratio accounting for rule-miss occurrences) of HTPDM remained 57.5 percent, whereas the worst-case precision ratio of WARD-HTP was 72.58 percent. In HTPDM, the gap between the worst and best instances is 7.5 percent, but in WARD-HTP, the difference is 2.44 percent. As a result, the disparity has shrunk by 5.08 percent. When adopting WARD-HTP, not only is the disparity in ratios decreased, but the best-case and worst-case relations are also increased (by 10% and 5.08%, respectively). WARD-HTP was fully implemented and tested against many input parameters to find the best configuration. We also experimented with different distance formulas, weighted training data methodologies (e.g., no weighted data), and training data year intervals. A Python implementation was used to test an 8-core Intel i7-4720 processor running at 3.35 GHz with 8 GB of RAM. We used HURDAT2 data for both training and testing. We fix not effort to infer points for storm paths for which no instructions were positively met, This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

A. Haversine DistanceFormula

In HTPDM, the Euclidean distance was compared to two distinct haversine distances: 300 km and 450 km (which is slightly less than the approximate average hurricane diameter). The model had undergone training. using data from 1950 to 2000 and then tested using data from 2001 to 2015. AprioriAll's parameters were set to minsup=3.75 percent and minconf=0.25. AprioriAll's training data was unweighted (i.e., weighting scale=1), and rule-miss extrapolation was turned off.The Euclidean distance choice has substantially lower accuracy ratios than the 300 km and 450 km haversine options, as seen in Table I. The 450-kilometer haversine distance yields the highest best-case and worst-case accuracy rates. While the ratio difference is smaller when using the Euclidean distance method, the total rise in both ratios makes the haversine distance of 450 kilometres the best choice. A 450-kilometer distance barrier allows for more flexibility in determining whether projections are right. Because 450 km corresponds to a size less than the typical storm width, such an increase in flexibility is permissible. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

Distfunction	Distthreshold	Best-caseratio	Worstcaseratio
Euclidean	1	36.28%	33.00%
Haversine	300	56.19%	51.20%
Haversine	**450**	**74.77%**	68.10%

TABLEI. RESULTS OF DIFFERENT DISTANCEFORMULAS

B. Weighted Training Data

Each WARD-HTP execution uses a different minsup and training data weighting approach, with a weighing scale of W=2.5. The absence of any weighing strategy is also an option in the weighting strategy options. This experiment compares three alternative weighting systems. The absence of data weighting is first investigated. After that, annual weighting is considered. W divided by the total number of years in the data Y (e.g., Y=50 years for training data from 1950 to 2000) raises the value of w in this technique. The index-weight technique, in which the weight is modified at each HDB entry,

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

is the third strategy used. This series of evaluations includes looks at different minimum support values. The three weighing algorithms are coupled against the minsup values equal to 2.5 percent, 3.75 percent, and 5% of the number of training data trajectories.While a greater minsup will result in fewer intriguing trajectory association rules, the rules that are produced may be more dependable for future predictions. Table II compares the number of frequent trajectories and trajectory association rules created with the best-case and worst-case hit rates for various minsup and weighing method pairings. Minconf=0.25 was used in all tests, coupled with HURDAT2 training data from 1950 to 2000, test data from 2001 to 2015, and haversine distance=450 km. WARD-HTP was fully implemented and tested against many input parameters to find the best configuration. We also experimented with different distance formulas, weighted training data methodologies (e.g., no weighted data), and training data year intervals. A Python implementation was used to test an 8-core Intel i7-4720 processor running at 3.35 GHz with 8 GB of RAM. We used HURDAT2 data for both training and testing.We do not attempt to extrapolate points for storm paths for which no rules were successfully met, Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

TABLEII. WEIGHING STRATEGIES AND *MINSUP* AGAINST RULES, PATTERNS AND CORRECTNESSRATIOS

	#rules	#patterns	Best-case hit rate	Worst-case hit rate
NoWeight				
minsup=2.5%	3,718	2,360	72.99%	69.75%
minsup=3.75%	635	544	74.77%	68.14%
minsup=5%	160	214	67.61%	57.25%
IndexWeight				
minsup=2.5%	14,228	7,112	75.00%	72.58%
minsup=3.75%	3,052	1,832	74.68%	71.38%
minsup=5%	466	429	75.22%	67.33%
YearWeight				
minsup=2.5%	25,797	11,899	**75.00%**	**72.58%**
minsup=3.75%	2,075	1,284	71.06%	67.37%
minsup=5%	769	619	71.68%	65.32%

When using data from 1950 to 2000, the best-case hit rate for WARD-HTP was calculated using a year-weighted method and minsup=2.5 percent. The hit rate as a result was 75.00 percent, with a worst-case scenario of 72.58 percent. Note that using minsup=5% and index-based weighting resulted in a slightly higher. Best-case hit rate of 75.22%,but a lower worst-case ratio of 67.33%.Because a 22% decline in Best-case hit rate for a 5.25% rise in worst-case best-case hit rate of 75.22 percent, but a lower worst-case ratio of 67.33 percent. Because a.22% decline in best-case hit rate for a 5.25 percent rise in worst-case hit rate is undoubtedly a "fair bargain," we choose 75 percent as the optimal generated data point. WARD-HTP was fully implemented and tested against many input parameters to find the best configuration. We also experimented with different distance formulas, weighted training data methodologies (e.g., no weighted data), and training data year intervals. A Python implementation was used to test an 8-core Intel i7-4720 processor running at 3.35 GHz with 8 GB of RAM. We used HURDAT2 data for both training and testing. We do not attempt to extrapolate points for storm paths for which no rules were successfully met, At minsup=2.5 percent, the year-based and index-based weighting procedures produced similar model results. We chose the year-based weighting strategy as the best solution since it fairly weights trajectory records throughout time. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

The year-based strategy counts all recordings from the most recent year at the maximum weight scale, whereas the index-based method only counts one record at the maximum.In this situation, both weighting approaches yielded similar outcomes, but the year-based strategy is preferred since it more closely matches the increase in support weight to real-world historical hurricane trajectory trends. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. This contains coordinate sequences, which may become uncommon in the future. AprioriAll's sequential mapping is done as a preparatory step to ensure. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

forecasting. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

A. **Year Interval for TrainingData**

The accuracy ratios generated by various training and test data intervals are also compared. Test data from 2001-2008 and 2001-2015 were compared to training data from 1900-2000 and 1950-2000. With minsup=2.5 percent, a yearly weighting approach, and haversine distance=450 km, all other factors are held constant. It's worth noting that HTPDM used data from 1900 to 2000 for training and 2001 to 2008 for testing.The correctness ratios (best-case and worst-case ratios) generated by all pairs of training and test data are shown in Table III. When all other parameters we regarded ideal throughout the preceding outcomes of this work are applied, we discover that the training data interval of 1950-2000 produces the best accuracy rates for test data from 2001-2015. To encourage analysis against more current hurricane patterns, the test data interval should always be 2001-2015, regardless of correctness ratio. The results reveal that no training-data/test-data pairing could produce accuracy ratios higher than those obtained by training the model using 1950-2000 data and testing it with 2001-2015 data.

Trainingdata	**1900-2000**		**1950-2000**	
Test data	Best-case	Worst-case	Best-case	Worst-case
2001-2015	74.16%	71.77%	**75.00%**	
	72.58%			
2001-2008	73.18%	70.62%	72.46%	
	69.93%			

TABLEIII. YEAR INTERVALS AGAINST CORRECTNESSRATIOS

VI Conclusion

The Weighted-Asset, Realistic-Distance Storm Trajectory Predictor (WARD-HTP) was introduced in this study as a new model for increasing the accuracy of existing data mining approaches to hurricane trajectory prediction. We used weighted training data from 1950 to 2000, as well as the haversine distance formula. The three contributions combine to generate WARD-HTP, which has a 75.0 percent correctness ratio for hurricane trajectory prediction, which is 10% greater than the present state-of-the-art. Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting. WARD-HTP was fully implemented and tested against many input parameters to find the best configuration. We also experimented with different distance formulas, weighted training data methodologies (e.g., no weighted data), and training data year intervals. A Python implementation was used to test an 8-core Intel i7-4720 processor running at 3.35 GHz with 8 GB of RAM. We used HURDAT2 data for both training and testing. We do not attempt to extrapolate points for storm paths for which no rules were successfully met, The correctness ratio of WARD-HTP is 72.58 percent when accounting for trajectories that could not be matched with any association rules, which is 15.08 percent higher than the worst case state-of-the-art ratio of 57.5 percent. WARD-HTP is now being extended to account for new characteristics such as air pressure distribution, water temperature, and wind direction. WARD-HTP was fully implemented and tested against many input parameters to find the best configuration. We also experimented with different distance formulas, weighted training data methodologies (e.g., no weighted data), and training data year intervals. A Python implementation was used to test an 8-core Intel i7-4720 processor running at 3.35 GHz with 8 GB of RAM. We used HURDAT2 data for both training and testing. We do not attempt to extrapolate points for storm paths for which no rules were successfully met, Property destruction, personal danger, and, in certain situations, widespread public disturbance are among the damages.[3]—inspire substantial research in the realm of storm forecasting.

References

[1] C. K. Leung, "Big data analysis and mining," in Encyclopedia of Information Science and Technology, 4th ed., pp.338-348.

[2] R. A. Pielke Jr., J. Gatz, C. W. Landsea, D. Collins, M. A. Saunders, & R. Musulin, "Normalized hurricane damage in the United States: 1900-2005," Natural Hazards Review, 9(1), Feb. 2008, pp.29-42.

[3] F. F. Townsend, The Federal Response to Hurricane Katrina: Lessons Learned,2006.

[4] Y. Su, S. Chelluboina, M. Hahsler, & M. H.

Dunham, "A new data mining model for hurricane intensity prediction," Proc. IEEE ICDM 2010 Workshops, pp.98-105.
[5] R. Yang, J. Tang, & D. Sun, "Association rule mining application for Atlantic tropical cyclone intensity changes," Weather and Forecasting, 26(3), June 2011, pp.337-353.
[6] O. S. Rozanova, J. Yu, & C. Hu, "Typhoon eye trajectory based on a mathematical model: comparing with observational data," Nonlinear Analysis:RealWorldApplications,11(3),June2010,pp.1847-1861.
[7] C. Roy & R. Kovordanyi, "Tropical cyclone track forecasting techniques: a review," Atmospheric Research, vol. 104-105, Feb. 2012, pp.40-69.
[8]H. Kim, J. Kim, C. Ho, & P. Chu, "Pattern classification of typhoon tracks using the fuzzy c-means clustering method," Journal of Climate, 24(2), Jan. 2011, pp.488-508
[9]P.Chandrasekar,K.Qian,H.Shahriar,& P.Bhattacharya, "Improving the prediction accuracy of decision tree mining with data preprocessing," Proc. IEEE COMPSAC 2017, vol. 2, pp. 481-484
[10] R. Agrawal & R. Srikant, "Mining sequential patterns," Proc. IEEE ICDE 1995, pp.3-14.
[11] X. Dong & D. C. Pi, "Novel method for hurricane trajectory prediction based on data mining," Natural Hazards and Earth System Sciences, vol. 13, 2013, pp.3211-3220.
[12]J. D. Hunter, "Matplotlib: a 2D graphics environment," Computing in Science and Engineering, 9(3), May-June 2007, pp.90-95.
[13] A. J. Kimberling, A. R. Buckley, P. C. Muehrcke, & J. O. Muehrcke, Map Use: Reading, Analysis, Interpretation, 7th ed.,2012.
[14] F. Chambat & B. Valette, "Mean radius, mass, and inertia for reference Earth models," Physics of the Earth and Planetary Interiors, 124(3-4), Aug. 2001, pp.237-253.
[15] C. W. Landsea, G. A. Vecchi, L. Bengtsson, & T. R. Knutson, "Impact of duration thresholds on Atlantic tropical cyclone counts," Journal of Climate, 23(10), May 2010, pp.2508-2519.