International Journal of Future Generation Communication and Networking
Vol.14, No. 1, (2021), pp. 4558- 4564

Tidal Wave Prediction Using Machine Learning

Mahika Godbole[1], Rucha Devshatwar[2], Pallavi Malekar[3], G.V. Raman Rao[*4], Vivek Saxena[+5]

Department of E&TC Engineering, SKNCOE, SPPU, Pune, India

[*]Scientist-D, CWPRS, Pune

[†]Research Assistant, CWPRS, Pune

[1]godbolemahika@gmail.com

[2]ruchadevshatwar@gmail.com

[3]pallavimalekar4520@gmail.com

[4]ramanarao_gv@cwprs.gov.in

[5]vivek.saxena.cwprs@gov.com

Abstract

Machine Learning is an application of artificial intelligence where a computer/machine learns from the past experiences (input data) and makes future predictions. The performance of such a system should be at least human level. Machine Learning is generally categorized into three types: Supervised Learning, Unsupervised Learning, Reinforcement learning.

Supervised Learning: In supervised learning the machine experiences the examples along with the labels or targets for each example. The labels in the data help the algorithm to correlate the features. Two of the most common supervised machine learning tasks are classification and regression. Unsupervised Learning: When we have unclassified and unlabelled data, the system attempts to uncover patterns from the data. There is no label or target given for the examples.

Reinforcement Learning: Reinforcement learning refers to goal-oriented algorithms, which learn how to attain a complex objective (goal) or maximize along a particular dimension over many steps. As machine learning technologies improve, they can be adapted to compile a continuous stream of real-time data collected locally with available forecasts into ever evolving and improving machine learning model parameters.

***Keyword**— Machine Learning, supervised learning, unsupervised learning, reinforcement learning.*

I. INTRODUCTION

Tides are the rise and fall of sea levels caused by the combined effects of the gravitational forces exerted by the Moon and the Sun and the rotation of the Earth. Most places in the ocean usually experience two high tides and two low tides each day (semi-diurnal tide), but some locations experience only one high and one low tide each day (diurnal tide).

The times and amplitude of the tides at the coast are influenced by the alignment of the Sun and Moon, by the pattern of tides in the deep ocean and by the shape of the coastline and near-shore bathymetry. Tides vary on timescales ranging from hours to years due to a number of factors, which determine the lunitidal interval. To make accurate records, tide gauges at fixed stations measure water level over time. Gauges ignore variations caused by waves with periods shorter than minutes. These data are compared to the reference (or datum) level usually called mean sea level.

- Tide changes proceed via the following stages:
- Sea level rises over several hours, covering the intertidal zone; flood tide.

ISSN: 2233-7853 IJFGCN

- The water rises to its highest level, reaching high tide.
- Sea level falls over several hours, revealing the intertidal zone; ebb tide.
- The water stops falling, reaching low tide.

While tides are usually the largest source of short-term sea-level fluctuations, sea levels are also subject to forces such as wind and barometric pressure changes, resulting in storm surges, especially in shallow seas and near coasts.

Tidal phenomena are not limited to the oceans but can occur in other systems whenever a gravitational field that varies in time and space is present. For example, the shape of the solid part of the Earth is affected slightly by Earth tide, though this is not as easily seen as the water tidal movements.

II. LITERATURE SURVEY

The tidal level variations are conventionally estimated using the harmonic analysis.[1] This paper presents an alternative method for prediction of tidal variations. Using this method it is possible to predict the tidal level variations at the reference station and fill the gap for the missing data. High value of of correlation coefficients obtained indicate the validity of technique for tidal level prediction.

Based on this study, an operational real time forecasting environment could be achieved when using a trained neural network. This technique can be conveniently used to generate missing data. However, if the training data are obtained from a different station, it should be ensured that the subordinate station is located in the nearby area say within the estuary or in a distance of about 20 - 30 km away and should not be located at a far distance of more than 100 km away.

Traditional methods of wave predictions have disadvantages of excessive data requirement, time consumption and are tedious to carry out. The present study makes use of a relatively new technique [2]of Artificial Neural Network which has been tried and tested in various coastal engineering applications.

In the present study [2] FFBP and NARX networks are used to predict waves at NMPT along the west coast of India. Predictions of waves at NMPT for one year have been carried out using yearlong wave data with FFBP network giving a satisfactory correlation coefficient ‘r’ value of 0.90 and 0.91 for the data set divided on monthly and weekly basis respectively.

Using NARX network prediction up to 25 weeks can be achieved with accuracy level greater than 0.94 using one week’s data and yearly prediction can be achieved with accuracy greater than 0.94 using one month’s data. Comparison of the results of FFBP network and NARX network showed NARX performing better than the later as the ‘r’ obtained in case of NARX was 0.94.

The review focused on the application of various soft computing techniques in predicting significant wave height [3]. The predictive efficiency of a machine learning approach depends upon quality and size of the data set available.

ANN takes more computational time and finds difficulty in determining the effective structure of the network. Hybrid models give better result than plain model. Wind speed, air temperature, sea surface temperature and wind direction has most to least influence in wave height prediction respectively

This paper has surveyed [4] the state-of-art literature in tidal analysis and forecasting methods for tsunami detection to identify publication trends, issues worthy of further investigation, and aspects that have not yet been completely exploited.

This paper is intended to provide [4] the communities of both researchers and practitioners with a broadly applicable, up to date coverage of tidal analysis and forecasting methodologies that have proven to be successful in a variety of circumstances, and that hold particular promise for success for the development or the improvement of Tsunami Alerting Devices for the detection of tsunamis events.

ISSN: 2233-7853 IJFGCN

In order to explore an effective method [5] to predict tidal level with a typhoon effect at Luchaogang, two tidal level prediction models, based on unique BP neural network and on a combination of BP neural network and KIA-based cubic B-spline curve, were introduced.

Then, using the data from Pudong Hydrological Bureau, experiments on the two models were conducted with six parameters of typhoon as the input parameters and corresponding tidal level data as the output parameters. The comparison results have showed that the BP-KIA model is superior to the unique BP neural network model.

This optimized model has not only improved the prediction accuracy of the highest tide level but also has accurate prediction results on time when the tidal level peak occurs. Through the BP-KIA prediction model, the trend of tide levels in 24 hours can be forecasted in real time. Based on the BP-KIA tidal level prediction model developed in this study, protective measures can be taken in advance to effectively reduce the impact of typhoon disasters.

III. Proposed System

A. *Data Collection*

Collection of data from CWPRS.

B. *Processing Data*

Next step is cleaning the data which includes rearranging the data and processing according to model and output requirements with dropping the unwanted rows and columns in the data.

C. *Training And Testing*

Third step is training and testing the data for predicting average Height of the tide. 70% data is trained in the model and 30%data is tested for Model predictions.

D. *Model Building And Deployment*

Actual model deployment included actual prediction by using data frame function as we are able to see the accuracy in our model with the less Error percentage Flask is a web application framework written in Python. It has multiple modules that make it easier for a web developer to write applications without having to worry about the details like protocol management, thread management, etc. Flask gives is a variety of choices for developing web applications and it gives us the necessary tools and libraries that allow us to build a web application. deploy model flask.

ISSN: 2233-7853 IJFGCN

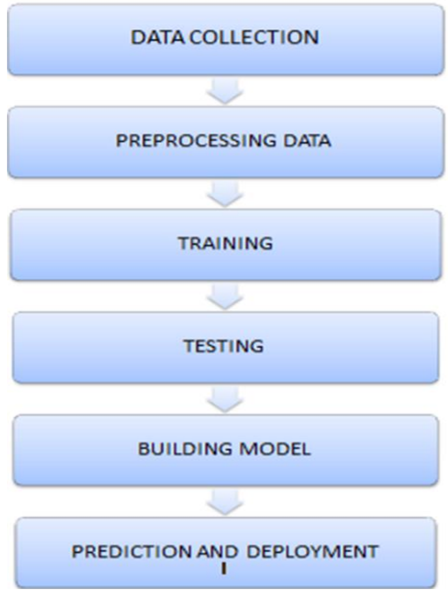


Fig. 1. Proposed System Flow

TABLE I
PARAMETERS

SR No.	Parameters	Description
1	DATA %	
2	WAVES	The number of waves
3	Hmax	Maximum Wave Height
4	Thm	Time Period of the wave
5	$H_{1/10}$	Average wave height of highest 1/10 of waves
6	$T_{1/10}$	Average wave period of highest 1/10 of waves
7	$H_{1/3}$	Average wave height of highest 1/3 of waves
8	$T_{1/3}$	Average wave period of highest 1/3 of waves
9	Havg	Average wave height waves
10	Tavg	Average wave period of waves
11	Eps	
12	Date(ddmmyy)	Date in Day/Month/Year

IV. RESULTS

ISSN: 2233-7853 IJFGCN

E. Data Import:

We took the raw data which is provided by CWPRS to us and then used pandas library to read the data as shown in Fig. 2

```
import pandas as pd
data1=pd.read_excel('april2013.xls')
data1.head()
```

	Wave Data of Minicoy for 04/04/2013	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	U
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
1	Time(mmddhhmm)	Data %	Waves	Hmax	Thm	H1/10	T1/10	H1/3	T1/3	Havg	
2	04040020.4RW	99.96745	227	45	11.16264	37.99153	9.350759	30.27074	7.740685	19.64317	
3	04040120.4RW	99.96745	241	50	11.25544	38.12	10.37978	29.58025	8.04792	18.43568	
4	04040220.4RW	99.96745	213	50	11.3664	41.16216	9.821616	32.56744	8.521933	20.55399	

Fig. 2. Data Import

F. Data Cleaning

Next step is cleaning the data which includes rearranging the data according to model and output requirements with dropping the unwanted rows and columns in data as shown in figure 3

```
import pandas as pd
data1=pd.read_excel('april2013.xls',skiprows=2)
data1.head()
```

	Time(mmddhhmm)	Data %	Waves	Hmax	Thm	H1/10	T1/10	H1/3	T1/3	Havg	Tavg	Ep
0	04040020.4RW	99.96745	227	45	11.16264	37.99153	9.350759	30.27074	7.740685	19.64317	5.252430	0.84218
1	04040120.4RW	99.96745	241	50	11.25544	38.12000	10.379780	29.58025	8.047920	18.43568	4.999222	0.83827
2	04040220.4RW	99.96745	213	50	11.36640	41.16216	9.821616	32.56744	8.521933	20.55399	5.609682	0.86327
3	04040320.4RW	99.96745	226	52	12.96285	39.78724	9.937702	31.00000	7.967224	19.74779	5.272195	0.84845
4	04040420.4RW	99.96745	193	65	12.11177	49.50495	11.167350	37.10769	9.651505	22.15026	6.177589	0.87655

Fig. 3. Data Cleaning

G. Training and Testing

Third step is training and testing the data for predicting average height of the tide. 70% data is trained in the model and 30%data is tested for model prediction.

```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
lr = LinearRegression()
X = data.drop(['Havg'],axis=1)
y = data['Havg']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3)
lr.fit(X_train, y_train)
lr.score(X_train, y_train),lr.score(X_test, y_test)
```

```
(0.9817946904014501, 0.9718675019398781)
```


ISSN: 2233-7853 IJFGCN

International Journal of Future Generation Communication and Networking
Vol.14, No. 1, (2021), pp. 4558- 4564

Fig. 4. Training and Testing

H. Model Results

Actual model deployment included actual prediction by using dataframe function as we are able to see the accuracy in our model with the less error percentage. Sample is shown in Fig. 5

```
df = pd.DataFrame(list(zip(y_test, y_pred)),
             columns =['y_test', 'y_pred'])
df.head(10)
```

	y_test	y_pred
0	17.95074	17.901803
1	24.93802	24.985422
2	29.06061	29.601772
3	31.18710	29.741836
4	25.20497	25.759990
5	18.43568	18.708849
6	30.34008	29.825012
7	32.67883	32.060529
8	26.65678	26.356574
9	27.85431	27.011256

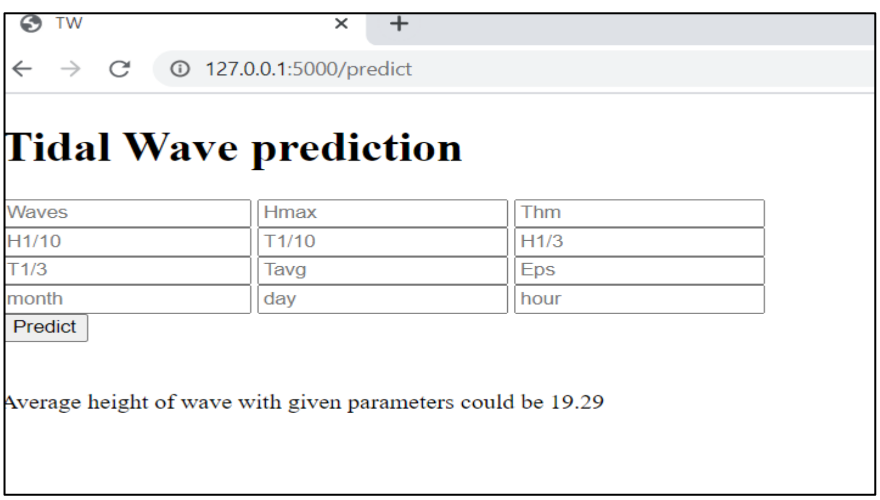
```
y_pred = lr.predict(X_test)
y_pred
```

```
array([17.90180346, 24.98542161, 29.60177165, 29.74183561, 25.75999001,
       18.70884882, 29.8250116 , 32.06052895, 26.35657425, 27.01125639,
       26.99622163, 29.37389025, 28.17757869, 23.40375363, 24.37179419,
       22.61940809, 24.77168904, 29.00557637, 28.718502  , 19.88151568,
       26.1647359 , 24.48977091])
```

Fig. 5. Model Result

I. Model Deployment

Actual model deployment results predicted by the model.



ISSN: 2233-7853 IJFGCN

Fig.6 Model Deployment Results

V. Conclusion

Prediction of tsunami and storms can be predicted with higher accuracy. The proposed system is helpful in assisting the ship movements at coastal environment. From proposed machine learning and artificial intelligence model we can predict the Tsunami movements change in the height of the wave coastal area harbour projects movements of ships these are things we can predict from the tidal wave prediction model. We can predict the height with utmost accuracy and this project will be helpful for the people living in the coastal regions. It will be helpful for their safety. The data generated by this model will be helpful for various coastal stations also.

References

[1] Meena, B.L. & Agrawal, Jagottam. (2015). Tidal level forecasting using ANN. Procedia Engineering. 116. 607-614. 10.1016/j.proeng.2015.08.332.

[2] Deepthi.I. Gopinath, G.S. Dwarakish, Wave Prediction Using Neural Networks at New Mangalore Port along West Coast of India, Aquatic Procedia, Volume 4,2015, Pages 143-150, ISSN 2214-241X

[3] Application of Machine Learning Techniques in Wave Height Forecasting In Marine Environment – A Review, Akhil C P & Paresh Chandra Deka, International Journal of Innovative Research in Science, Engineering and Technology (An ISO 3297: 2007 Certified Organization) Website: www.ijirset.com ,Vol. 6, Issue 5, May 2017

[4] Consoli, S., D. R. Recupero and Vanni Zavarella. “A survey on tidal analysis and forecasting methods for Tsunami detection.” ArXiv abs/1403.0135 (2014): n. pag. Submitted on December 2013

[5] Wenjuan Wang, Hongchun Yuan, "A Tidal Level Prediction Approach Based on BP Neural Network and Cubic B-Spline Curve with Knot Insertion Algorithm", Mathematical Problems in Engineering, vol. 2018, Article ID 9835079, 9 pages, 2018.

[6] Consoli, S., D. R. Recupero, and V. Zavarella. "A survey on tidal analysis and forecasting methods for Tsunami detection. arXiv 2014." arXiv preprint arXiv:1403.0135.

[7] Shuto, Nobuo, and Koji Fujima. "A short history of tsunami research and countermeasures in Japan." Proceedings of the Japan Academy, Series B 85, no. 8, pp. 267-275, 2009.

[8] Beltrami, Gian Mario. "An ANN algorithm for automatic, real-time tsunami detection in deep-sea level measurements." Ocean Engineering 35, no. 5-6, 572-587, 2008.

[9] Bishop, Christopher M. Neural networks for pattern recognition. Oxford university press, 1995.

[10] Foreman, Michael George Garvin. Manual for tidal heights analysis and prediction. Institute of Ocean Sciences, Patricia Bay, 1979.

ISSN: 2233-7853 IJFGCN