**Summary of the research paper as an outcome of PBL-I project**

**ON**

**“Demographic Location Surface Water Mapping”**

A PBL I report submitted in partial fulfillment of the requirements for the degree of Bachelor of Technology in Computer Science & Engineering

BACHELOR OF TECHNOLOGY in COMPUTER SCIENCE & ENGINEERING

Submitted By

Abhishek Rajput 22070122007

Arnav Jain 22070122030

Atul Goyal 22070122039

Janmejay Pandya 22070122086

UNDER THE GUIDANCE OF

Dr. Nilima Zade

Designation



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**Symbiosis Institute of Technology, Pune**

**Symbiosis International (Deemed University)**

**CERTIFICATE**

This is to certify that the PBL I Project work entitled “Demographic Location Surface Water Mapping” is carried out by Abhishek Rajput, Arnav Jain, Atul Goyal, Janmejay Pandya**,** in partial fulfillment for the award of the degree of **Bachelor of Technology** in **Computer Science & Engineering**, Symbiosis Institute of Technology Pune, Symbiosis International (Deemed University) Pune, India during the academic year 2023-2024.

|  |  |  |
| --- | --- | --- |
| -------------- |  | -------------- |
| Name of Co-Guide |  | Name of Guide |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**UNDERTAKING**

We undertake that we have prepared a paper as per the following details during our PBL-I project. The paper has been submitted to the project guide.

**Tentative title of the of the paper:** Demographic Location Surface Water Mapping

**Status of the paper publication:** Draft prepared

Abhishek Rajput 22070122007 Signature

Arnav Jain 22070122030 Signature

Atul Goyal 22070122039 Signature

Janmejay Pandya 22070122086 Signature

|  |  |  |
| --- | --- | --- |
| -------------- |  | -------------- |
| Name & Signature of Co-Guide |  | Name & Signature of the Guide |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

## TABLE OF CONTENTS

|  |  |  |
| --- | --- | --- |
|  | | **Page** |
| **Certificate** | | i |
| **Undertaking** | | ii |
| **Table of Contents** | | iii |
|  | |  |
| 1 | Problem statement of the project | 1 |
| 2 | Abstract of the Paper | 1 |
| 3 | Summary of the Literature Review | 2 |
| 4 | Brief description of the methodology | 3 |
| 5 | Summary of results | 3 |

**Problem statement of the project**

To develop an effective satellite image classification system Utilizing the Image Processing technique and ML based algorithms while addressing challenges posed by sun and mountain shadows. With shadow correction methods, we aim to accurately classify land and water areas in satellite imagery for a specific time period, considering the impact of shadows on image interpretation.

**Abstract of the Paper**

The "Demographic Location Surface Water Mapping" project introduces a novel method to precisely categorise land and water regions in satellite images of the demographic region, while tackling the difficulties caused by sun and mountain shadows. The main method used is the Modified Normalised Difference Water Index (MNDWI), which is specifically intended to improve the clarity of water features. Nevertheless, shadows can impede precise identification of bodies of water. In order to address this problem, the project includes powerful shadow correction algorithms that correct the influence of shadows on the satellite photos.

The objective of the project is to create a satellite image classification system that combines MNDWI with shadow correction techniques. This integration will enhance the accuracy and dependability of monitoring and analysing surface water bodies. The utilisation of data in this technique has substantial consequences for the management of water resources and urban development in the Pune region. Policymakers can acquire useful information to make educated decisions about water distribution, conservation, and long-term sustainability by closely observing changes in water bodies over time.

The report provides a comprehensive description of the methodology employed in the Project, which encompasses data acquisition, image preprocessing, MNDWI application, binary image conversion, and time series analysis. The report outlines the utilisation of various AI & ML models, including Linear Regression, Polynomial Regression, Random Forest, Gradient Boosting, Support Vector Machine XGBoost and Neural Network. The findings illustrate how the initiative might contribute to sustainable decision-making by effectively managing urban growth while simultaneously preserving the environment.

**Summary of the Literature Review**

D. Naik etal (2021)., "An Implementation of Satellite Image Classification and Analysis using Machine Learning with ISRO LISS IV," Int. J. Comput. Electron. T. Jain etal (2022) "Satellite Image Classification and Analysis using Machine Learning with ISRO LISS IV," International Research Journal of Engineering and Technology (IRJET). S. Singh etal (2023), "Environmental monitoring with machine learning," EPRA Int. J. Multidiscip. Reddy, etal (2020). Satellite Image Classification for Environmental Analysis using Image Processing. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(3), 1513-1517. Ouchra etal (2023). Machine Learning Algorithms for Satellite Image Classification Using Google Earth Engine and Landsat Satellite Data: Morocco Case Study. S. M. Moorthi etal (2011) "Kernel based learning approach for satellite image classification using support vector machine," in Proc. IEEE Recent Advances in Intelligent Computational Systems (RAICS). Simoes etal(2021). Satellite Image Time Series Analysis for Big Earth Observation Data. Du etal (2016)."Water bodies’ mapping from Sentinel 2 imagery with modified normalized difference water index at 10 m spatial resolution produced by sharpening the SWIR band." Remote Sensing. Feng etal(2015)."Urban flood mapping based on unmanned aerial vehicle remote sensing and random forest classifier—A case of Yuyao, China." Water. Pekel etal (2016)."High resolution mapping of global surface water and its long term changes." Nature, 540(7633), 418 422. Rushikesh Kulkarni etal (2022) . Detecting, extracting, and mapping of inland surface water using Landsat 8 Operational Land Imager. Santini etal (2010)."A two step optimization procedure for assessing water constituent concentrations by hyperspectral remote sensing techniques: An application to the highly turbid Venice lagoon waters. Masek etal (2006). A Landsat surface reflectance dataset for North America, 1990–2000. IEEE Geoscience and Remote Sensing Letters. M. Rodell et al (2018), "Emerging trends in global freshwater availability.

**Brief description of the methodology**

We acquire satellite imagery of the Pune region from reputable sources such as the United States Geological Survey (USGS). In order to accurately depict the landscape, we make adjustments for the effects of sunlight and shadows cast by mountains in specific bands, such as band 3 and band 7. We utilise the Modified Normalised Difference Water Index (MNDWI) on corrected images. This rating improves visibility by accentuating distinctions between water and land characteristics. We transform the retrieved photos into a format where land pixels are assigned a value of 0 and water pixels are assigned a value of 1. This process streamlines data in order to facilitate analysis. In order to examine the temporal variations in surface water in Pune, we extract the land and water areas from the image on a regular basis. This enables the application of time series analysis. Utilising machine learning algorithms to classify land and water areas for future predictions, followed by the development of a frontend to visually present the images.

**Summary of results**

The study presents a comparative analysis of different machine learning models, evaluating their performance indicators, including RMSE (Root Mean Square Error) and R² (Coefficient of Determination), for predicting both water and land areas.

The Random Forest and Gradient Boosting models demonstrated superior performance in terms of both RMSE and R², suggesting enhanced predictive accuracy and fit for both water and land surfaces in comparison to other models.

Linear Regression and Polynomial Regression demonstrated stable performance across various degrees, however they were less effective compared to more intricate models.

The performance of Support Vector Machine and K-Nearest Neighbours was satisfactory.

The performance of XGBoost on this dataset was subpar, as indicated by the notably low R² scores, which imply a lack of adequate fit.

The performance of the Neural Networks was severely subpar, either as a result of overfitting or incorrect configuration.