**Mini Project Report on**



**LAND USE LAND COVER ANALYSIS USING MACHINE LEARNING AND REMOTE SENSING**



**Submitted in partial fulfillment of the requirement for the award of the degree of**

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE & ENGINEERING**

**Submitted by:**

**Kanuri Pranay** **2019417**

***Under the Mentorship of***

**Dr. Hemant Singh Pokhariya**

**Professor**



**Department of Computer Science and Engineering**

**Graphic Era (Deemed to be University)**

**Dehradun, Uttarakhand**

**January 2024**



**CANDIDATE’S DECLARATION**

I hereby certify that the work which is being presented in the project report entitled **Land Use Land Cover Analysis Using Machine Learning And Remote Sensing** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineeringof the Graphic Era (Deemed to be University), Dehradun shall be carried out by the under the mentorship of **Dr. Hemant Singh Pokhariya, Professor**, Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Kanuri Pranay 2019417 **signature**

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Description** | **Page No.** |
| Chapter 1 | Introduction | **1-2** |
| Chapter 2 | Literature Survey | **3-4** |
| Chapter 3 | Methodology | **5-6** |
| Chapter 4 | Result and Discussion | **7** |
| Chapter 5 | Conclusion and Future Work | **8** |
|  | References | **9** |

**Chapter 1**

**Introduction**

Land Use and Land Cover (LULC) analysis is a crucial tool that enables us to gain insights into how people utilize the land and what types of coverage exist on the Earth's surface. This analytical approach is instrumental in the sustainable management of land and resources. In our current project, we leverage advanced technology and utilize data obtained from space to enhance our understanding of various aspects related to land. By harnessing the capabilities of remote sensing and space-based data, we aim to acquire valuable information that contributes to more informed decision-making and responsible land management practices.

In our project, the primary focus is on Land Use Land Cover (LULC) data, providing valuable insights into the various elements present on the Earth's surface, including trees, buildings, and other features. This information is crucial for understanding the composition and spatial distribution of different land uses.

Additionally, we are incorporating Land Surface Temperature (LST) data into our analysis. LST helps us assess how hot or cold the land surface is. This temperature is influenced by the land cover types, such as the presence of trees or urban areas, as well as weather conditions. By integrating LULC and LST data, we aim to gain a comprehensive understanding of the Earth's surface dynamics and environmental conditions, facilitating more informed decision-making in land management and resource planning.

The code initiates by utilizing diverse Earth observation data, focusing on analyzing rainfall patterns by integrating precipitation data with vibrant visuals. Simultaneously, it delves into the study of Land Surface Temperature (LST), generating visual displays to elucidate the variations in land temperature, highlighting areas that experience heat or cold.

Additionally, the code incorporates population data, employing distinct colors to depict areas where people reside. Furthermore, elevation data is harnessed to illustrate the topographical variations of the land, employing a spectrum of colors to smoothly represent the ups and downs of the terrain. Through this multifaceted approach, the code provides a comprehensive and visually engaging representation of Earth's environmental dynamics, encompassing precipitation, temperature, population distribution, and topography.

The central concept revolves around amalgamating various layers—such as precipitation patterns, Land Surface Temperature (LST), population distribution, and elevation data—into a single composite image referred to as 'combined.' This consolidated image provides a holistic depiction of the environmental conditions and dynamics.

Transitioning to machine learning, the code employs a sophisticated tool known as Random Forest Regression. This intelligent tool operates with a meticulously selected subset of data, ensuring effective learning. The dataset is bifurcated into two segments—one for training the model and the other for assessing its accuracy. Through this approach, the code leverages

The trained model utilizes factors such as rainfall, elevation, and population distribution to make predictions about Land Surface Temperature (LST). These predictions are then translated into a specialized 'LST Prediction' map, offering a visual representation of the anticipated temperatures across different areas.

To validate the accuracy of the predictions, a scatter chart is employed. This chart, characterized by lines and dots, serves as a comparative tool, illustrating how closely the model comprehends the relationship between environmental variables and Land Surface Temperature. Through this evaluation process, the scatter chart provides insights into the reliability and effectiveness of the model in capturing the complex interactions that influence temperature variations in the environment.  
Ultimately, the insights derived from the scatter chart contribute significantly to our understanding of how human activities influence changes in the land and how the environment responds. This comprehensive analysis aids in making informed decisions about sustainable land management practices, fostering a deeper commitment to caring for our planet. By integrating advanced technologies, Earth observation data, and machine learning, this approach not only enhances our comprehension of environmental dynamics but also provides valuable guidance for responsible and impactful decision-making to ensure the well-being of our planet.

**Chapter 2**

**Literature Survey**

The comprehension of land-use/land cover change has undergone a transition from simplicity to a more realistic and complex understanding in recent decades. Initially, studies primarily focused on the physical aspects of change. However, as the global environmental change research agenda gained momentum, scientists recognized the profound influence of land surface processes on climate due to land use/cover change. In the mid-1970s, it became evident that land cover change, particularly modifications in surface albedo, significantly impacts surface-atmosphere energy exchanges, thereby influencing regional climate (Otterman, 1974; Charney and Stone, 1975; Sagan et al., 1979).This realization prompted a broader exploration of the impacts of land-use/cover change on ecosystems, goods, and services. Notable concerns include the worldwide impact on biotic diversity (Sala et al., 2000), soil degradation (Trimble and Crosson, 2000), and the ability of biological systems to meet human needs (Vitousek, 1997; Praveen, B. 2017).

Throughout history, humans have modified land to secure essentials for survival. However, the contemporary rate of exploitation far surpasses historical levels, leading to unprecedented changes in ecosystems and environmental processes at local, regional, and global scales. Current land use/land cover changes now encapsulate a spectrum of environmental concerns for the global human population, including climate change, biodiversity depletion, and pollution of water, soil, and air. Consequently, monitoring and mitigating the adverse consequences of land use/land cover change while ensuring the sustainable production of essential resources have become paramount priorities for researchers and policymakers worldwide (Erle and Pontius, 2007).

The escalating impact of unsustainable human activities is emerging as a key environmental concern, particularly in deteriorating water quality. Understanding the intricate relationship between land use and water quality is instrumental in identifying threats to river water quality (Ding et al., 2015). Additionally, gaining insights into the 'access' to sanitation is crucial for human survival, emphasizing the interconnectedness of land use changes with broader environmental and public health challenges (Parveen et al., 2015; Praveen et al., 2017).

The literature on land use/land cover change studies provides a comprehensive exploration of how human populations have transformed the terrestrial biosphere over more than 8000 years, resulting in the creation of anthropogenic biomes (Ellis, 2011). Recent interest in land use/land cover (LULC) change spans a diverse group of researchers, ranging from those modeling spatio-temporal patterns to those investigating the causes, impacts, and consequences (Verburg et al., 1999; Brown et al., 2000; Theobald, 2001).

Central to these studies is the dynamic relationship between land use and land cover. While changes in land cover by land use do not necessarily imply land degradation, shifting land use patterns, often due to social causes, can lead to consequential changes in land cover. These alterations have wide-ranging effects on biodiversity, water resources, radiation budgets, and other processes, collectively influencing climate and the biosphere (Riebsame et al., 1994).

The literature underscores the complex interplay of human-environment interactions, emphasizing the need for comprehensive approaches that consider the multifaceted influences on land use and cover changes. Researchers employ various methods, from modeling techniques to investigating socio-economic drivers, contributing to a nuanced understanding of the evolving landscapes shaped by human activities over time.

Land use/land cover studies employing remote sensing and GIS techniques play a crucial role in optimizing land utilization. Access to information on existing land use and land cover is essential for making informed decisions about land management. Additionally, the capability to monitor the dynamic changes in land use resulting from both the evolving demands of a growing population and natural forces shaping the landscape is indispensable.

Land is in a constant state of transformation due to various natural and man-made processes, making it imperative to study spatio-temporal patterns of intra and inter-urban forms. Understanding the evolution of urban systems remains a primary objective in urban research. This knowledge is vital for staying abreast of changes and updating land cover maps, enabling effective natural resource management (Xiaomei and Rong Qing, 1999).

Remote sensing devices capture responses based on various characteristics of the land surface, encompassing both natural and artificial features. To interpret this data, analysts rely on elements such as tone, texture, pattern, shape, size, shadow, site, and association to derive information about land cover. The process of generating remotely sensed data or images involves using different types of sensors deployed on various platforms at varying heights above the terrain and at different times of the day and year.

Developing a simple classification system for this diverse range of remotely sensed data is challenging. There's a consensus that no single classification can be universally applied to all types of imagery and at all scales. In response to this complexity, Anderson's pioneering work in 1976 led to the creation of a general-purpose classification scheme compatible with remote sensing data. This scheme, often referred to as the United States Geological Survey (USGS) classification scheme, has played a crucial role in facilitating consistent and meaningful interpretation of remotely sensed information.

Top of Form

Since the launch of the first remote sensing satellite, Landsat-1, in 1972, land use/land cover studies have been conducted on various scales to cater to different user needs. An example of such studies is the waste land mapping of India, which was performed on a 1:1 million scale by the National Remote Sensing Agency (NRSA) using Landsat multispectral scanner data from 1980-82. This study estimated that approximately 16.2% of the land in India was categorized as waste land based on the provided data.Over time, a series of studies have indicated that Landsat Thematic Mapper is well-suited for providing general extensive synoptic coverage of large areas. This has led to a reduction in the necessity for expensive and time-consuming ground surveys, as Landsat data has proven valuable for validation purposes. The utilization of satellite imagery, especially from platforms like Landsat, has enhanced the efficiency and cost-effectiveness of land use/land cover studies on a global scale.

**Chapter 3**

**Methodology**

**3.1 Data Collection:**

* Collect Earth observation datasets, including precipitation, Land Surface Temperature (LST), population, and elevation data.

**3.2 Precipitation Analysis:**

* Sum precipitation data to reveal rainfall patterns.
* Rename the result as 'rain' and visualize it on the map with a color palette representing different precipitation levels.

**3.3 Land Surface Temperature (LST) Processing:**

* Calculate the mean of LST, convert it to Celsius, and rename it as 'LST.'
* Visualize LST on the map using a color palette to represent temperature variations.

**3.4 Population Data Integration:**

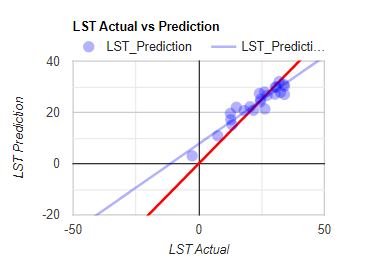
* Filter population data based on the study area boundaries.
* Sort the data based on the temporal aspect and select the latest available snapshot.
* Rename the dataset as 'pop' and visualize it on the map with a color palette representing population density.
  1. **Elevation Data Visualization:**
* Select the 'elevation' band from the nasadem dataset.
* Visualize elevation on the map using a color palette to represent different elevation levels.
  1. **Data Fusion:**
* Combine the processed datasets (rain, LST, pop, elevation) into a single image named 'combined.'
* Apply a mask using population data to focus on areas with human presence.
  1. **Data Sampling:**
* Sample 10,000 pixels within the defined region (geometry) at a scale of 100 meters.
* Add a random column for stratified sampling and visualize the sampled points on the map.
  1. **Data Splitting:**
* Split the sampled data into training (80%) and testing (20%) sets based on the random column.
  1. **Random Forest Regression:**
* Implement Random Forest Regression using 50 trees.
* Train the model using the training set with LST as the dependent variable and 'rain,' 'elevation,' and 'pop' as independent variables.
* Set the output mode to regression and print the model details.
  1. **Prediction Mapping:**
* Apply the trained regression model to the 'combined' image to create a spatially explicit 'LST Prediction' layer.
* Visualize the prediction on the map using a color palette.
  1. **Accuracy Assessment:**
* Classify the test data using the trained regression model for accuracy testing.
* Create a scatter chart comparing actual LST values with predicted values, showing the relationship and model accuracy.
* Customize the chart with data opacity, axis titles, series representation, and trendlines for clarity.
  1. **Results Presentation:**
* Display the accuracy chart and print the sizes of the training and testing sets.
* Evaluate the effectiveness of the Random Forest Regression model in predicting Land Surface Temperature based on the selected variables.

**Chapter 4**

**Result and Discussion**

The result obtained was satisfactory as the model was able to predict the Land Surface Temperature rightfully with a good level of accuracy. We obtained a machine-learned module that can be used to predict the precipitation using the Random Forest module.

The accuracy of the above model was tested on area includes the India, Bangladesh, Nepal, Bhutan, Parts of Pakistan, Sri Lanka, Myanmar Dataset using the Google Search Engine. Below is the comparison graph obtained providing the LST (Actual) vs its predicted value.



**Figure 4.1** LST Actual vs Prediction

**Chapter 5**

**Conclusion and Future Work**

This report was an attempt to determine and integrate diverse Earth observation datasets to gain insights into environmental dynamics in India. The comprehensive analysis included precipitation, Land Surface Temperature (LST), population, and elevation data. We obtained the prediction of LST with the help of Machine Learning that is, Random Forest Regression Technique. The implementation of Random Forest Regression, utilizing 50 trees and considering 'rain,' 'elevation,' and 'pop' as predictors for LST, demonstrated a promising approach for predicting Land Surface Temperature.

The model can be extended and refined in several ways. The inclusion of additional environmental variables, such as land cover and vegetation indices, could enhance the model's predictive capabilities. Moreover, exploring alternative machine learning algorithms and parameter tuning may further improve the accuracy of LST predictions.

Additionally, we can expand the geographical scope of the study to cover different parts of the world.

**References**

1. Adeniyi, P. O., & Omojola, A. (1999). Landuse/landcover change evaluation in Sokoto-Rima basin of NW Nigeria based on archival remote sensing and GIS techniques. Geoinformation Technology Applications of Remote Sensing Applications for Resource and Environmental Management in Africa. Ed. by Adeniyi, PO: African Association of Remote Sensing of the Environment.

2. Begum, N., Narayana, J., & SL, A. K. (2010). Land Use/Land Cover Changes in the Catchment of Water Bodies in and Around Davangere City, Karnataka. International Journal of Ecology and Environmental Sciences, 36(4), 277-280.

3. Bisht, B. S., & Kothyari, B. P. (2001). Land-cover change analysis of Garur Ganga watershed using

GIS/remote sensing technique. Journal of the Indian Society of Remote Sensing, 29(3), 137-141.

4. Brown, K., McIlveen, H., & Strugnell, C. (2000). Nutritional awareness and food preferences of young consumers. Nutrition & Food Science, 30(5), 230-235.

5. Bushra. Praveen.(2017). Monitoring and Assesment of Landslide from Agastamuni to Sonprayag. International Journal of Recent Research in Social Sciences and Humanities (IJRRSSH), 4(2), pp. 92-118.

6. Charney, J., Stone, P. H., & Quirk, W. J. (1975). Drought in the Sahara: a biogeophysical feedback mechanism.

science, 187(4175), 434-435.

7. Dhinwa, P. S., Pathan, S. K., Sastry, S. V. C., Rao, M., Majumder, K. L., Chotani, M. L., ... & Sinha, R. L. P.

(1992). Land use change analysis of Bharatpur district using GIS. Journal of the Indian society of Remote Sensing, 20(4), 237-250.

8. Ding, J., Jiang, J., Fu, L., Liu, Q., Peng, Q. & Kang, M. 2015. Impacts of Land Use on Surface Water Quality in a Subtropical River Basin: A Case Study of the Dongjiang River Basin, Southeastern China. Journal Of Water, 7, 4427-4445.

9. Erle, E., & Pontius, R. (2007). Land-use and land-cover change. Encyclopaedia of Earth.(Eds.). Cutler J. Cleveland (Washington, DC: Environmental Information Coalition, National Council for Science and the Environment). Last Retrieved January, 19, 2008.

10. Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., ... & Helkowski, J. H. (2005).

Global consequences of land use. science, 309(5734), 570-574.

11. Javed, A., & Khan, I. (2012). Land use/land cover change due to mining activities in Singrauli industrial belt, Madhya Pradesh using remote sensing and GIS. Journal of Environmental Research And Development Vol, 6(3A).

12. Khan. J.H, Parveen. S, Ahmed. N. Regional Analysis of Sanitation Facilities in Uttar Pradesh. Journal of Humanities and Social Sciences, (IOSR-JHSS), 20(10), pp. 48-56.