

1 DENGUEWATCH: A SYSTEM FOR REAL-TIME
2 DENGUE MONITORING AND FORECASTING IN ILOILO
3 PROVINCE

4 A Special Problem Proposal
5 Presented to
6 the Faculty of the Division of Physical Sciences and Mathematics
7 College of Arts and Sciences
8 University of the Philippines Visayas
9 Miag-ao, Iloilo

10 In Partial Fulfillment
11 of the Requirements for the Degree of
12 Bachelor of Science in Computer Science by

13 AMODIA, Kurt Matthew A.
14 BULAONG, Glen Andrew C.
15 ELIPAN, Carl Benedict L.

16 Francis D. DIMZON
17 Adviser

18 April 28, 2025

Abstract

Dengue fever remains a significant public health concern in the Philippines, with cases rising dramatically in recent years. Nationwide outbreaks have placed immense strain on healthcare systems, underscoring the need for innovative approaches to surveillance and response. In Iloilo City, this national trend was reflected in a significant surge, with the Iloilo Provincial Health Office reporting 4,585 cases and 10 fatalities as of August 10, 2023—a 319% increase from the previous year’s 1,095 cases and one death. This study developed a centralized system for monitoring and modernizing data management of dengue cases in public health institutions, making it more efficient and acceptable. Using data gathered from the Iloilo Provincial Health Office and online sources, several deep learning models were trained to predict dengue cases, utilizing weather variables and historical case data as inputs. These models included Long Short-Term Memory (LSTM), ARIMA, Seasonal ARIMA, Kalman Filter (KF), and a hybrid KF-LSTM model. The models underwent time series cross-validation strategies to mimic real-world conditions as closely as possible and were evaluated using metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). The LSTM model demonstrated the best performance with the lowest RMSE of 16.90, followed by the hybrid KF-LSTM model at 25.56. The LSTM model was then integrated into the system to provide forecasting features that could support health institutions by offering actionable insights for proactive intervention strategies.

Keywords: ARIMA, artificial intelligence, dengue prediction, LSTM, Kalman Filter, deep learning, climate variables, public health, outbreak mitigation

Contents

43	1 Introduction	1
44	1.1 Overview	1
45	1.2 Problem Statement	2
46	1.3 Research Objectives	2
47	1.3.1 General Objective	2
48	1.3.2 Specific Objectives	2
49	1.4 Scope and Limitations of the Research	3
50	1.5 Significance of the Research	4
51	2 Review of Related Literature	6
52	2.1 Dengue	6
53	2.2 Outbreak Definition	6
54	2.3 Existing System: RabDash DC	7
55	2.4 Deep Learning	8
56	2.5 Kalman Filter	8
57	2.6 Weather Data	9
58	2.7 Chapter Summary	9

59	3 Research Methodology	10
60	3.1 Research Activities	11
61	3.1.1 Gather Dengue Data and Climate Data to Create a Com-	
62	plete Dataset for Forecasting	11
63	3.1.2 Develop and Evaluate Deep Learning Models for Dengue	
64	Case Forecasting	13
65	3.1.3 Integrate the Predictive Model into a Web-Based Data An-	
66	alytics Dashboard	17
67	3.1.4 System Development Framework	17
68	3.2 Development Tools	19
69	3.2.1 Software	19
70	3.2.2 Hardware	20
71	3.2.3 Packages	20
72	3.3 Application Requirements	22
73	3.3.1 Backend Requirements	22
74	3.3.2 User Interface Requirements	23
75	3.3.3 Security and Validation Requirements	25
76	3.4 Calendar of Activities	26
77	4 Results and Discussion/System Prototype	27
78	4.1 Data Gathering	27
79	4.2 Exploratory Data Analysis	28
80	4.3 Outbreak Detection	32
81	4.4 Model Training Results	33
82	4.4.1 LSTM Model	33

83	4.4.2	ARIMA Model	35
84	4.4.3	Seasonal ARIMA (SARIMA) Model	36
85	4.4.4	Kalman Filter Model	37
86	4.5	Model Simulation	39
87	4.6	System Prototype	39
88	4.6.1	Home Page	39
89	4.6.2	User Registration, Login, and Authentication	40
90	4.6.3	Encoder Interface	42
91	4.6.4	Admin Interface	50
92	4.7	User Testing	55
93		References	57
94		A Appendix Title	60
95		B Resource Persons	61

96 List of Figures

97	3.1	Workflow for forecasting the number of weekly dengue cases . . .	10
98	3.2	Testing Process for DengueWatch	18
99	3.3	Entity-Relationship Database Schema Hybrid Diagram for DengueDash	
100		Database Structure	22
101	3.4	Use Case Diagram for Admins	23
102	3.5	Use Case Diagram for Encoder	24
103	4.1	Snippet of the Combined Dataset	28
104	4.2	Data Contents	28
105	4.3	Dataset Statistics	29
106	4.4	Distribution of Dengue Cases	29
107	4.5	Frequency of Dengue Cases in Different Intervals	30
108	4.6	Trend of Dengue Cases	31
109	4.7	Ranking of Correlations	31
110	4.8	Ranking of Correlations (with lagged effects)	32
111	4.9	Training Folds - Window Size 5	34
112	4.10	Testing Folds - Window Size 5	34
113	4.11	ARIMA Prediction Results for Test Set	35

114	4.12 Seasonal ARIMA Prediction Results for Test Set	36
115	4.13 Kalman Filter Prediction Results for Test Set	38
116	4.14 Predicted vs Actual Dengue Cases 2025	39
117	4.15 Home Page	40
118	4.16 Sign Up Page	41
119	4.17 Login Page	41
120	4.18 First Part of Case Report Form	42
121	4.19 Second Part of Case Report Form	43
122	4.20 Bulk Upload of Cases using CSV	44
123	4.21 Dengue Reports	45
124	4.22 Detailed Case Report	46
125	4.23 Update Report Dialog	47
126	4.24 Delete Report Alert Dialog	48
127	4.25 Forecasting Page	49
128	4.26 Retraining Configurations	50
129	4.27 Start Retraining	51
130	4.28 Retraining Results	51
131	4.29 List of Verified Accounts	52
132	4.30 List of Pending Accounts	53
133	4.31 Account Details	53
134	4.32 DRU Registration	54
135	4.33 List of DRUs	54
136	4.34 DRU details	55

137	A.1 LSTM Prediction Results for Test Set	60
-----	--	----

138 **List of Tables**

139	3.1	Timetable of Activities for 2024	26
140	3.2	Timetable of Activities for 2025	26
141	4.1	Comparison of different models for dengue prediction	33
142	4.2	Comparison of Window Sizes	33
143	4.3	Comparison of SARIMA performance for each fold	37
144	4.4	Comparison of KF-LSTM performance for each fold	38
145	4.5	Computed System Usability Scores per Participant	55

Chapter 1

Introduction

1.1 Overview

From 2020 to 2022, dengue cases declined due to reduced surveillance during the COVID-19 pandemic (WHO, 2023), but cases surged in 2023 as restrictions were lifted. This year saw an increase in dengue outbreaks worldwide, with over five million cases and more than 5,000 deaths reported in over 80 countries (Bosano, 2023). Dengue is endemic in the Philippines, leading to longer and more widespread seasonal outbreaks. Globally, dengue infections have increased significantly, posing a major public health challenge. The World Health Organization reported a ten-fold rise in cases between 2000 and 2019, with a peak in 2019 when the disease spread across 129 countries (WHO, 2024).

Iloilo City and Province are intensifying efforts to curb the rising dengue cases (Lena, 2024). As of August 10, 2023, the Iloilo Provincial Health Office recorded 4,585 cases and 10 deaths, a 319% increase from last year's 1,095 cases and one death. Governor Arthur Defensor Jr. confirmed that the province has reached the dengue outbreak threshold based on Department of Health (DOH). Local government units (LGUs) have been informed, and the province's disaster management office is on blue alert, indicating disaster mode. (Perla, 2024)

In Iloilo City, 649 dengue cases were recorded this year 2024, with two deaths. Cases cluster in 40 out of 180 barangays, meaning multiple cases are being reported in these areas over several weeks. The city's health officer, Dr. Roland Jay Fortuna, reported high utilization of non-COVID-19 hospital beds, reaching over 76%, prompting concerns about hospital capacity.

170 This study explores the monitoring and forecasting of dengue outbreaks by an-
171 alyzing key factors such as temperature, relative humidity, and historical dengue
172 cases, using different models. The findings aim to provide an advanced, AI-driven
173 alternative for dengue prevention and control, targeting agencies like the Depart-
174 ment of Health (DOH). By aligning with the national AI Roadmap, particularly
175 in Iloilo City, this research aspires to improve outbreak responses through cutting-
176 edge technology rather than traditional reporting methods.

177 1.2 Problem Statement

178 Dengue remains a critical public health challenge worldwide, with cases increasing
179 due to the easing of COVID-19 restrictions and heightened global mobility. While
180 a temporary decline in cases was observed during the pandemic (2020–2022) due
181 to reduced surveillance efforts, 2023 marked a resurgence, with over five million
182 cases and more than 5,000 deaths reported across 80 countries. In dengue-endemic
183 regions like the Philippines, the threat is particularly severe. In Iloilo City and
184 Province, dengue cases rose by 319% as of August 2023, overwhelming local health-
185 care systems. This surge strained resources, with over 76% of non-COVID-19 hos-
186 pital beds occupied by dengue patients, highlighting the urgent need for effective
187 predictive tools. The lack of a reliable system to monitor and forecast dengue
188 outbreaks contributes to delayed interventions, exacerbating public health risks
189 and healthcare burdens in the region.

190 1.3 Research Objectives

191 1.3.1 General Objective

192 This study aims to develop an AI-based dengue forecasting and monitoring system
193 for Iloilo City and Province. The researchers will train and compare multiple deep
194 learning models to predict dengue case trends based on climate data and historical
195 dengue cases to help public health officials in possible dengue case outbreaks.

196 1.3.2 Specific Objectives

197 Specifically, this study aims to:

- 198 1. Gather dengue data from the Iloilo Provincial Health Office and climate data
199 (including temperature, rainfall, wind, and humidity) from online sources.
200 Combine and aggregate these data into a unified dataset to facilitate com-
201 prehensive dengue case forecasting.
- 202 2. Evaluate deep learning models for predicting dengue cases using metrics
203 such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE),
204 and Mean Squared Error (MSE). Compare the performance of these models
205 to determine the most accurate forecasting approach.
- 206 3. Develop a web-based analytics dashboard that integrates a predictive model
207 and provides data management system for dengue cases in Iloilo City and
208 the Province.
- 209 4. Assess the usability and effectiveness of the analytics dashboard through
210 structured feedback and surveys involving health professionals and policy-
211 makers.

212 1.4 Scope and Limitations of the Research

213 This study aims to gather dengue data from the Iloilo Provincial Health Office
214 and climate data from online sources such as PAGASA or weatherandclimate.com.
215 These data will be preprocessed, cleaned, and combined into a unified dataset to
216 facilitate comprehensive dengue case forecasting. However, the study is limited by
217 the availability and completeness of historical data. Inconsistent or missing data
218 points may introduce biases and reduce the quality of predictions. Furthermore,
219 the granularity of the data will be in a weekly format.

220 To evaluate deep learning models for predicting dengue cases, the study will
221 train and compare the performance of various models, using metrics like Mean
222 Absolute Error (MAE) and Root Mean Square Error (RMSE). While these models
223 aim to provide accurate forecasts, their performance is heavily influenced by the
224 quality and size of the dataset. Limited or low-quality data may lead to suboptimal
225 predictions. Additionally, the models cannot fully account for external factors
226 such as public health interventions or socio-economic conditions which may impact
227 dengue transmission dynamics.

228 The study also involves developing a web-based analytics dashboard that in-
229 tegrates predictive models and provides a data management system for dengue
230 cases in Iloilo City and the Province. This dashboard will offer public health
231 officials an interactive interface to visualize dengue trends, input new data, and

232 identify risk areas. However, its usability depends on feedback from stakeholders,
233 which may vary based on their familiarity with analytics tools. Moreover, exter-
234 nal factors such as limited internet connectivity or device availability in remote
235 areas may affect the system’s adoption and effectiveness. While the dashboard
236 provides valuable insights, it cannot incorporate all factors influencing dengue
237 transmission, emphasizing the need for ongoing validation and refinement.

238 1.5 Significance of the Research

239 This study’s development of an AI-based dengue forecasting and monitoring sys-
240 tem has wide-reaching significance for various stakeholders in Iloilo City:

- 241 • **Public Health Agencies:** Organizations like the Department of Health (DOH)
242 and local health units in Iloilo City and Province stand to benefit greatly
243 from the system. With dengue predictions, we can help these agencies opti-
244 mize their response strategies and implement targeted prevention measures
245 in high-risk areas before cases escalate.
- 246 • **Local Government Units (LGUs):** LGUs can use the system to support
247 their disaster management and health initiatives by proactively addressing
248 dengue outbreaks. The predictive insights allow for more efficient planning
249 and resource deployment in barangays and communities most vulnerable to
250 outbreaks, improving overall public health outcomes.
- 251 • **Healthcare Facilities:** Hospitals and clinics, which currently face high bed
252 occupancy rates during dengue season will benefit from early outbreak fore-
253 casts that can help in managing patient inflow and ensuring adequate hos-
254 pital capacity.
- 255 • **Researchers and Policymakers:** This AI-driven approach contributes valu-
256 able insights for researchers studying infectious disease patterns and policy-
257 makers focused on strengthening the national AI Roadmap. The system’s
258 data can support broader initiatives for sustainable health infrastructure
259 and inform policy decisions on resource allocation for dengue control.
- 260 • **Community Members:** By reducing the frequency and severity of outbreaks,
261 this study ultimately benefits the community at large. This allows for timely

262 awareness campaigns and community engagement initiatives, empowering
263 residents with knowledge and preventative measures to protect themselves
264 and reduce the spread of dengue.

Chapter 2

Review of Related Literature

2.1 Dengue

Dengue disease is a tropical and subtropical mosquito-borne viral illness and is a major health concern in the Philippines (Bravo, Roque, Brett, Dizon, & L’Azou, 2014). The majority of individuals with dengue experience no symptoms. Fever is the most common symptom, typically 4 to 7 days after being bitten by an infected mosquito (Zhou & Malani, 2024). In recent years, the trend of dengue cases in the Philippines has shown notable fluctuations, with periodic outbreaks occurring every 3 to 5 years, often influenced by climatic and environmental changes. According to the Department of Health (DOH), the number of reported cases has steadily increased over the past decades, attributed to urbanization, population growth, and inadequate vector control measures (World Health Organization (WHO), 2018). Moreover, studies suggest that El Niño and La Niña events have significant effects on dengue incidence, with warmer temperatures and increased rainfall providing favorable breeding conditions for mosquitoes (Watts, Burke, Harrison, Whitmire, & Nisalak, 2020). The study of Carvajal et. al. highlights the temporal pattern of dengue cases in Metropolitan Manila and emphasizes the significance of relative humidity as a key meteorological factor, alongside rainfall and temperature, in influencing this pattern (Carvajal et al., 2018).

2.2 Outbreak Definition

The definition of an outbreak is a critical factor in disease surveillance, as it determines the threshold at which an unusual increase in cases is considered a

288 public health concern. Studies suggest that outbreak thresholds should be context-
289 specific, given the variability in transmission dynamics across different locations
290 (Runge-Ranzinger et al., 2016). Outbreak definitions defined using the Endemic
291 Channel often base thresholds on 2 standard deviations (SD) above the mean
292 number of historic dengue cases. Other studies (Hemisphere, 2015) also used an
293 alert threshold of a 5-year mean plus 2 standard deviations. A study by (Brady,
294 Smith, Scott, & Hay, 2015) discussed that definitions of dengue outbreaks differ
295 significantly across regions and time, making them inconsistent and incomparable.

296 From the studies above, this research implements an outbreak definition basing
297 the threshold on 2 standard deviations (SD) above the mean number of historic
298 dengue cases. It is important to note that for future outbreak definitions, addi-
299 tional local context like available hospital space etc., must be taken into account
300 for a more effective outbreak definition.

301 2.3 Existing System: RabDash DC

302 RabDash, developed by the University of the Philippines Mindanao, is a web-
303 based dashboard for rabies data analytics. It combines predictive modeling with
304 genomic data, enabling local health authorities to optimize interventions and al-
305 locate resources more effectively. RabDash’s modules include trend visualization,
306 geographic hotspot mapping, and predictive forecasting, utilizing Long Short-
307 Term Memory (LSTM) models for time-series forecasting (RabDashDC, 2024).

308 For DengueWatch, RabDash serves as a strong inspiration, particularly in
309 its monitoring, historical trend visualization, and forecasting capabilities. These
310 features align well with the needs of dengue control efforts, providing real-time
311 insights into outbreak trends and enabling more effective, data-driven decision-
312 making. RabDash’s architecture is relevant to the DengueDash, as dengue out-
313 breaks similarly require time-series forecasting models. By using LSTM, RabDash
314 effectively models trends in outbreak data, which provides a framework for adapt-
315 ing LSTM to dengue forecasting. Research indicates that LSTM models outper-
316 form traditional methods, such as ARIMA and MLP, in handling the complexities
317 of time-dependent epidemiological data (Ligue & Ligue, 2022).

318 2.4 Deep Learning

319 The study of (Ligue & Ligue, 2022) highlights how data-driven models can help
320 predict dengue outbreaks. The authors compared traditional statistical meth-
321 ods, such as non-seasonal and seasonal autoregressive integrated moving average
322 (ARIMA), and traditional feed-forward network approach using a multilayer per-
323 ceptron (MLP) model with a deep learning approach using the long short-term
324 memory (LSTM) architecture in their prediction model. They found that the
325 LSTM model performs better in terms of accuracy. The LSTM model achieved a
326 much lower root mean square error (RMSE) compared to both MLP and ARIMA
327 models, proving its ability to capture complex patterns in time-series data (Ligue
328 & Ligue, 2022). This superior performance is attributed to LSTM’s capacity
329 to capture complex, time-dependent relationships within the data, such as those
330 between temperature, rainfall, humidity, and mosquito populations, all of which
331 contribute to dengue incidence (Ligue & Ligue, 2022).

332 2.5 Kalman Filter

333 The Kalman Filter is another powerful tool for time-series forecasting that can be
334 integrated into our analysis. It provides a recursive solution to estimating the state
335 of a linear dynamic system from a series of noisy measurements. Its application
336 in epidemiological modeling can enhance prediction accuracy by accounting for
337 uncertainties in the data (Li et al., 2022). Studies have shown that Kalman filters
338 are effective in predicting infectious disease outbreaks by refining estimates based
339 on observed data. A study published in *Frontiers in Physics* utilized the Kalman
340 filter to predict COVID-19 deaths in Ceará, Brazil. They found that the Kalman
341 filter effectively tracked the progression of deaths and cases, providing critical in-
342 sights for public health decision-making (Ahmadini et al., 2021). Another research
343 article in *PLOS ONE* focused on tracking the effective reproduction number (R_t)
344 of COVID-19 using a Kalman filter. This method estimated the growth rate of
345 new infections from noisy data, demonstrating that the Kalman filter could main-
346 tain accurate estimates even when case reporting was inconsistent (Arroyo-Marioli,
347 Bullano, Kucinskas, & Rondón-Moreno, 2021).

348 Our study will compare ARIMA, seasonal ARIMA, Kalman Filter, and LSTM
349 models using our own collected dengue case data along with weather data to
350 identify the most effective model for real-time forecasting.

351 2.6 Weather Data

352 The relationship between weather patterns and mosquito-borne diseases is inher-
353 ently nonlinear, meaning that fluctuations in disease cases do not respond propor-
354 tionally to changes in climate variables(Colón-González, Fezzi, Lake, & Hunter,
355 2013) Weather data, such as minimum temperature and accumulated rainfall, are
356 strongly linked to dengue case fluctuations, with effects observed after several
357 weeks due to mosquito breeding and virus incubation cycles. Integrating these
358 lagged weather effects into predictive models can improve early warning systems
359 for dengue control(Cheong, Burkart, Leitão, & Lakes, 2013). A study also suggests
360 that weather-based forecasting models using variables like mean temperature and
361 cumulative rainfall can provide early warnings of dengue outbreaks with high sen-
362 sitivity and specificity, enabling predictions up to 16 weeks in advance(Hii, Zhu,
363 Ng, Ng, & Rocklöv, 2012).

364 We will utilize weather data, including variables such as temperature, rainfall,
365 and humidity, as inputs for our dengue forecasting model. Given the strong, non-
366 linear relationship between climate patterns and dengue incidence, these weather
367 variables, along with their lagged effects, are essential for enhancing prediction
368 accuracy and providing timely early warnings for dengue outbreaks.

369 2.7 Chapter Summary

370 This chapter reviewed key literature relevant to our study, focusing on existing
371 systems, predictive modeling techniques and the role of weather data in forecast-
372 ing dengue outbreaks. We examined systems like RabDash DC, which integrates
373 predictive modeling with real-time data to inform public health decisions, provid-
374 ing a foundational structure for our Dengue Watch System. Additionally, deep
375 learning approaches, particularly Long Short-Term Memory (LSTM) networks,
376 were highlighted for their effectiveness in time-series forecasting, while alternative
377 methods such as ARIMA and Kalman Filters were considered for their ability to
378 model complex temporal patterns and handle noisy data.

379 The literature further underscores the significance of weather variables—such
380 as temperature and rainfall—in forecasting dengue cases. Studies demonstrate
381 that these variables contribute to accurate outbreak prediction models. Lever-
382 aging these insights, our study will incorporate both weather data and historical
383 dengue case counts to build a reliable forecasting model.

Chapter 3

Research Methodology

This chapter lists and discusses the specific steps and activities that will be performed to accomplish the project. The discussion covers the activities from pre-proposal to Final SP Writing.

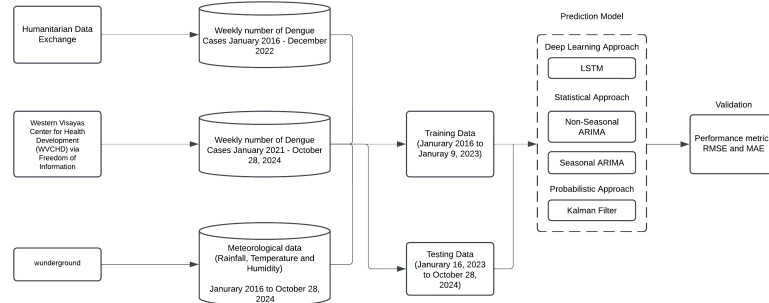


Figure 3.1: Workflow for forecasting the number of weekly dengue cases

This summarizes the workflow for forecasting the number of weekly dengue cases. This workflow focuses on using statistical, deep learning, and probabilistic models to forecast the number of reported dengue cases. The approach involves deploying several models for prediction, including ARIMA and Seasonal ARIMA as statistical approaches, LSTM as a deep learning approach, and the Kalman Filter as a probabilistic approach. These methods are compared with each other to determine the most accurate model.

3.1 Research Activities

3.1.1 Gather Dengue Data and Climate Data to Create a Complete Dataset for Forecasting

Acquisition of Dengue Case Data

The historical dengue case dataset used in this study was obtained from the Humanitarian Data Exchange and the Western Visayas Center for Health Development (WVCHD) via Freedom of Information (FOI) requests. The decision to use weekly intervals was driven by the need for precision and timeliness in capturing fluctuations in dengue cases and weather conditions. Dengue transmission is influenced by short-term changes in weather variables such as rainfall and temperature, which impact mosquito breeding and virus transmission cycles. A weekly granularity allowed the model to better capture these short-term trends, enabling more accurate predictions and responsive public health interventions.

Moreover, using a weekly interval provided more data points for training the models compared to a monthly format. This is particularly critical in time series modeling, where larger datasets help improve the robustness of the model and its ability to generalize to new data. Also, the collection of weather data was done by utilizing web scraping techniques to extract weekly weather data (e.g., rainfall, temperature, and humidity) from Weather Underground (wunderground.com).

Data Fields

- **Time.** Represents the specific year and week corresponding to each entry in the dataset.
- **Rainfall.** Denotes the observed average rainfall, measured in millimeters, for a specific week.
- **Humidity.** Refers to the observed average relative humidity, expressed as a percentage, for a specific week.
- **Max Temperature.** Represents the observed maximum temperature, measured in degrees Celsius, for a specific week.
- **Average Temperature.** Represents the observed average temperature, measured in degrees Celsius, for a specific week.

- 427 • **Min Temperature.** Represents the observed minimum temperature, mea-
428 sured in degrees Celsius, for a specific week.
- 429 • **Wind.** Represents the observed wind speed, measured in miles per hour
430 (mph), for a specific week.
- 431 • **Cases.** Refers to the number of reported dengue cases during a specific
432 week.

433 **Data Integration and Preprocessing**

434 The dengue case data was integrated with the weather data to create a com-
435 prehensive dataset, aligning the data based on corresponding timeframes. The
436 dataset underwent a cleaning process to address any missing values, outliers, and
437 inconsistencies to ensure its accuracy and reliability. To ensure that all features
438 and the target variable were on the same scale, a MinMaxScaler was applied to
439 normalize both the input features (climate data) and the target variable (dengue
440 cases).

441 **Exploratory Data Analysis (EDA)**

- 442 • Analyzed trends, seasonality, and correlations between dengue cases and
443 weather factors.
- 444 • Created visualizations like time series plots and scatterplots to highlight
445 relationships and patterns in the data.

446 **Outbreak Detection**

447 To detect outbreaks, we computed the outbreak threshold value of dengue cases
448 using the formula,

$$\text{Outbreak Threshold Value} = \mu + 2\sigma \quad (3.1)$$

449 where μ is the historical mean and σ is the standard deviation.

450 3.1.2 Develop and Evaluate Deep Learning Models for 451 Dengue Case Forecasting

452 The deep learning models were developed and trained to forecast weekly dengue
453 cases using historical weather data (rainfall, temperature, wind, and humidity)
454 and dengue case counts. The dataset was normalized and divided into training and
455 testing sets, ensuring temporal continuity to avoid data leakage. The methodology
456 for preparing and training the model are outlined below.

457 Data Preprocessing

458 The raw dataset included weekly aggregated weather variables (rainfall, tempera-
459 ture, wind, humidity) and dengue case counts. The "Time" column was converted
460 to a datetime format to ensure proper temporal indexing. To standardize the data
461 for training, MinMaxScaler was employed, normalizing the feature values and tar-
462 get variable to a range of 0 to 1. This step ensured that the models could efficiently
463 process the data without being biased by feature scaling differences.

464 LSTM Model

465 The dataset was split into training and test sets to evaluate the model's perfor-
466 mance and generalizability:

- 467 • **Training Set:** 80% of the data (572 sequences) was used for model training,
468 enabling the LSTM to learn underlying patterns in historical dengue case
469 trends and their relationship with weather variables.
- 470 • **Test Set:** The remaining 20% of the data (148 sequences) was reserved for
471 testing

472 To prepare the data for LSTM, a sliding window approach was utilized. Se-
473 quences of weeks of normalized features were constructed as input, while the
474 dengue case count for the subsequent week was set as the target variable. This
475 approach ensured that the model leveraged temporal dependencies in the data for
476 forecasting. To enhance the performance of the LSTM model in predicting dengue
477 cases, Bayesian Optimization was employed using the Keras Tuner library. The
478 tuning process aimed to minimize the validation loss (mean squared error) by
479 adjusting key model hyper-parameters. The search space is summarized below:

480 **LSTM units:**

- 481 • min value: 32
- 482 • max value: 128
- 483 • step: 16
- 484 • sampling: linear

485 **Learning Rate:**

- 486 • min value: 0.0001
- 487 • max value: 0.01
- 488 • step: None
- 489 • sampling: log

490 The tuner was instantiated with:

- 491 • **max trials = 10:** Limiting the search to 10 different configurations
- 492 • **executions per trial = 3:** Running each configuration thrice to reduce
493 variance
- 494 • **validation split = 0.2:** Reserving 20% of the training data for validation

495 The hyperparameter tuning was conducted for three different window sizes of
496 data: 5, 10, and 20. This allows the model to have the optimal hyperparameters
497 used for each window size. Training was conducted over 100 epochs with early
498 stopping to prevent overfitting while maintaining computational efficiency. A
499 batch size of 1 was used, enabling the model to process individual sequences,
500 which is suitable for smaller datasets but results in longer training times. The
501 Adam optimizer, known for its adaptive learning capabilities and stability was
502 employed.

503 To validate the effectiveness of the model, cross-validation was implemented.
504 However, standard k-fold cross-validation randomly shuffles the data, which isn't
505 suitable for time series since the order of observations is important. To address
506 this, a time series-specific cross-validation strategy was used with TimeSeriesS-
507 plit from the scikit-learn library. This method creates multiple train-test splits

508 where each training set expands over time and each test set follows sequentially.
509 This approach preserves the temporal structure of the data while helping reduce
510 overfitting by validating the model across different time segments.

511 After training, predictions on both the training and test datasets were rescaled
512 to their original scale using the inverse transformation of MinMaxScaler. Model
513 performance was evaluated using the mean squared error (MSE), root mean
514 squared error (RMSE) and mean absolute error (MAE).

515 **ARIMA**

516 The ARIMA model was utilized to forecast weekly dengue cases, leveraging histor-
517 ical weather data—including rainfall, maximum temperature, and humidity—as
518 exogenous variables alongside historical dengue case counts as the primary depen-
519 dent variable. The dataset was partitioned into training (80%) and testing (20%)
520 sets while maintaining temporal consistency.

521 To identify the optimal ARIMA configuration, a comprehensive grid search
522 was performed across the following parameter ranges:

- 523 • Autoregressive order (p): 0 to 3
- 524 • Differencing order (d): 0 to 2
- 525 • Moving average order (q): 0 to 3

526 Each combination of (p,d,q) was used to fit an ARIMA model, and perfor-
527 mance was evaluated based on the mean squared error (MSE) between the pre-
528 dicted and actual dengue cases on the test set. The parameter set that achieved
529 the lowest MSE was selected as the final model configuration.

530 Following model selection, the best-fit ARIMA model was retrained on the
531 training set and subsequently used to forecast dengue cases for the test period.
532 The predictions were assigned to the **PredictedCases** column in the test dataset.
533 Model performance was further assessed using key evaluation metrics, including
534 MSE, root mean squared error (RMSE), and mean absolute error (MAE). Visual
535 comparisons between actual and predicted dengue cases were produced through
536 line plots to better illustrate the model’s forecasting accuracy.

537 Seasonal ARIMA (SARIMA)

538 The SARIMA modeling process began with data preprocessing, which included
539 handling missing values through interpolation or imputation, and standardizing
540 features to ensure stable model training. The dataset was then split into training
541 and testing sets in an 80:20 ratio, preserving the temporal order of observations.
542 Seasonality analysis was conducted using time series decomposition and autocor-
543 relation plots, which revealed a periodicity of 52 weeks—justifying the adoption
544 of a seasonal model. To fine-tune the model, a grid search was performed over a
545 range of SARIMA parameters $(p,d,q)(P,D,Q)[S]$, while stationarity was validated
546 using the Augmented Dickey-Fuller (ADF) test. The model was then trained
547 on the dataset using rainfall, temperature, and humidity as exogenous variables,
548 with convergence ensured by setting a maximum number of iterations. Residual
549 diagnostics were used to confirm that residuals were uncorrelated, indicating a
550 good model fit. For evaluation, forecasts were compared against actual values,
551 and results were visualized with line plots. Finally, to validate the model’s gener-
552 alizability across different time periods, Time Series Cross-Validation with three
553 folds was applied. This allowed assessment of the model’s performance on multi-
554 ple time segments, providing insights into its robustness in real-world forecasting
555 scenarios.

556 Kalman Filter:

- 557 • Input Variables: The target variable (Cases) was modeled using three re-
558 gressors: rainfall, max temperature, and humidity.
- 559 • Training and Testing Split: The dataset was split into 80% training and
560 20% testing to evaluate model performance.
- 561 • Observation Matrix: The Kalman Filter requires an observation matrix,
562 which was constructed by adding an intercept (column of ones) to the re-
563 gressors.

564 The Kalman Filter’s Expectation-Maximization (EM) method was employed
565 for training, iteratively estimating model parameters over 10 iterations. After
566 training, the smoothing method was used to compute the refined state estimates
567 across the training data. Observation matrices for the test data were constructed
568 in the same manner as for the training set, ensuring compatibility with the learned
569 model parameters. On the test data, the Kalman Filter applied these parameters
570 to predict and correct the estimated dengue cases, providing more stable and
571 accurate forecasts compared to direct regression models. Additionally, a hybrid

572 Kalman Filter–LSTM (KF-LSTM) model was developed to combine the strengths
573 of both approaches. In this setup, the LSTM model was first used to predict
574 dengue cases based on historical data and weather features. The Kalman Filter
575 was then applied as a post-processing step to the LSTM predictions, smoothing
576 out noise and correcting potential errors.

577 **3.1.3 Integrate the Predictive Model into a Web-Based** 578 **Data Analytics Dashboard**

579 **Dashboard Design and Development**

- 580 • Design an intuitive, user-friendly web-based dashboard incorporating:
 - 581 – Interactive visualizations of yearly dengue case trends.
 - 582 – Data input and update forms for dengue and weather data.
 - 583 – Map display of dengue cases in each district in Iloilo City

584 **Model Integration and Deployment**

- 585 • Deploy the best-performing model within the dashboard as a backend service
586 to enable real-time or periodic forecasting.

587 **3.1.4 System Development Framework**

588 The Agile Model is the birthchild of both iterative and incremental approaches
589 in Software Engineering. It aims to be flexible and effective at the same time by
590 being adaptable to change. It's also important to note that small teams looking
591 to construct and develop projects quickly can benefit from this kind of method-
592 ology. As the Agile Method focuses on continuous testing, quality assurance is a
593 guarantee since bugs and errors are quickly identified and patched.

594 **Design and Development**

595 After brainstorming and researching the most appropriate type of application to
596 accommodate both the prospected users and the proposed solutions, the team
597 has decided to proceed with a web application. Given the time constraints and
598 available resources, it has been decided that the said means is the most pragmatic

599 and practical move. The next step is to select modern and stable frameworks
600 that align with the fundamental ideas learned by the researchers in the university.
601 The template obtained from WVCHD and Iloilo Provincial Epidemiology and
602 Surveillance Unit was meticulously analyzed to create use cases and develop a
603 preliminary well-structured database that adheres to the requirements needed
604 to produce a quality application. The said use cases serve as the basis of general
605 features. Part by part, these are converted into code, and with the help of selected
606 libraries and packages, it resulted in the desired outcome that may still modified
607 and extended to achieve scalability.

608 Testing and Integration

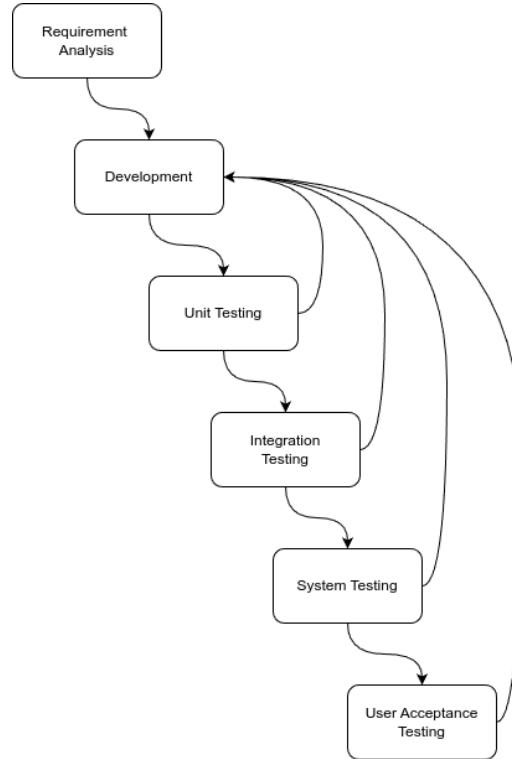


Figure 3.2: Testing Process for DengueWatch

609 Implementing testing is important to validate the system's performance and ef-
610 ficacy. Thus a series of tests were conducted to identify and resolve bugs during
611 the developmental phase. Each feature was rigorously tested to ensure quality as-
612 surance, with particular emphasis on prerequisite features, as development cannot
613 progress properly if these fail. Because of this, integration between each feature

614 serves as a pillar for a cohesive user experience. Since dengue reports include
615 confidential information, anonymized historical dengue reports were used to train
616 the model and create the foundational architecture of the system. By using func-
617 tional tests, data validation and visualization can be ensured for further continual
618 improvements. Security testing is also important as it is needed to safeguard
619 confidential information when the system is deployed. It includes proper authen-
620 tication, permission views, and mitigating common injection attacks. Finally, a
621 user acceptance test from the prospected users, in this case, doctors, nurses, and
622 other health workers is crucial to assess its performance and user experience. It
623 enables the developers to confirm if the system meets the needs of the problem.

624 **3.2 Development Tools**

625 **3.2.1 Software**

626 **Github**

627 GitHub is a cloud-based platform that tracks file changes using Git, an open-
628 source version control system (*About GitHub and Git - GitHub Docs*, n.d.). It is
629 used in the project to store the application’s source code, manage the system’s
630 source version control, and serve as a repository for the Latex files used in the
631 actual research.

632 **Visual Studio Code**

633 Visual Studio Code is a free, lightweight, and cross-platform source code editor
634 developed by Microsoft (*Why Visual Studio Code?*, 2021). As VS Code supports
635 this project’s programming and scripting languages, it was chosen as the primary
636 source code editor.

637 **Django**

638 Django is a free and open-sourced Python-based web framework that offers an
639 abstraction to develop and maintain a secure web application. As this research
640 aims to create a well-developed and maintainable application, it is in the best
641 interest to follow an architectural pattern that developers and contributors in the

642 future can understand. Since Django adheres to Model-View-Template (MVT)
643 that promotes a clean codebase by separating data models, business logic, and
644 presentation layers, it became the primary candidate for the application’s back-
645 bone.

646 **Next.js**

647 A report by Statista (2024) claims that React is the most popular front-end frame-
648 work among web developers. However, React has limitations that can be a nui-
649 sance in rapid software development, which includes routing and performance op-
650 timizations. This is where Next.js comes in—a framework built on top of React.
651 It offers solutions for React’s deficiency, making it a rising star in the framework
652 race.

653 **Postman**

654 As the application heavily relies on the Application Programming Interface (API)
655 being thrown by the backend, it is a must to use a development tool that facilitates
656 the development and testing of the API. Postman is a freemium API platform
657 that offers a user-friendly interface to create and manage API requests (*What is*
658 *Postman? Postman API Platform*, n.d.).

659 **3.2.2 Hardware**

660 The web application was developed on laptop computers with minimum specifica-
661 tions of an 11th-generation Intel i5 CPU and 16 gigabytes of RAM. Furthermore,
662 an RTX 3060 mobile GPU was also utilized to retrain the integrated model locally.

663 **3.2.3 Packages**

664 **Django REST Framework**

665 Django Rest Framework (DRF) is a third-party package for Django that provides a
666 comprehensive suite of features to simplify the development of robust and scalable
667 Web APIs (Christie, n.d.). These services include Serialization, Authentication
668 and Permissions, Viewsets and Routers, and a browsable API .

669 Leaflet

670 One of the features of the web application is the ability to map the number
671 of cases using a Choropleth Map. Leaflet is the only free, open-sourced, and
672 most importantly, stable JavaScript package that can do the job. With its ultra-
673 lightweight size, it offers a comprehensive set of features that does not trade
674 off performance and usability (*Leaflet — an open-source JavaScript library for*
675 *interactive maps*, n.d.).

676 Chart.js

677 Another feature of the application is to provide users with informative, approach-
678 able data storytelling that is easy for everyone to understand. The transformation
679 of pure data points and statistics into figures such as charts is a big factor. Thus,
680 there is a need for a package that can handle this feature without compromising
681 the performance of the application. Chart.js is a free and open-source JavaScript
682 package that is made to meet this criteria as it supports various types of charts
683 (*Chart.js*, n.d.).

684 Tailwind CSS

685 Using plain CSS in production-quality applications can be counterproductive.
686 Therefore, CSS frameworks were developed to promote consistency and accelerate
687 the rapid development of web applications (Joel, 2021). One of these is Tailwind,
688 which offers low-level utility classes that can be applied directly to each HTML
689 element to create a custom design (*Tailwind CSS - Rapidly build modern web-*
690 *sites without ever leaving your HTML.*, n.d.). Given the limited timeline for this
691 project, using this framework is a wise choice due to its stability and popularity
692 among developers.

693 Shadcn

694 Shadcn offers a collection of open-source UI boilerplate components that can be
695 directly copied and pasted into one's project. With the flexibility of the provided
696 components, Shadcn allows developers to have full control over customization and
697 styling. Since this is built on top of Tailwind CSS and Radix UI, it is supported
698 by most modern frontend frameworks, including Next.js (Shadcn, n.d.).

699 **Zod**

700 Data validation is integral in this web application since it will handle crucial data
701 that will be used for analytical inferences and observations. Since Zod is primarily
702 used for validating and parsing data, it ensures proper communication between
703 the client and the server (Zod, n.d.).

704 **3.3 Application Requirements**

705 **3.3.1 Backend Requirements**

706 **Database Structure Design**

707 Determining how data flows and how it would be structured is crucial in creating
708 the system as it defines how extendible and flexible it would be for future features
709 and updates. Thus, creating a comprehensive map of data ensures proper nor-
710 malization that eliminates data redundancy and improves data integrity. Figure
711 3.3 depicts the designed database schema that showcases the relationship between
712 the application’s entities.

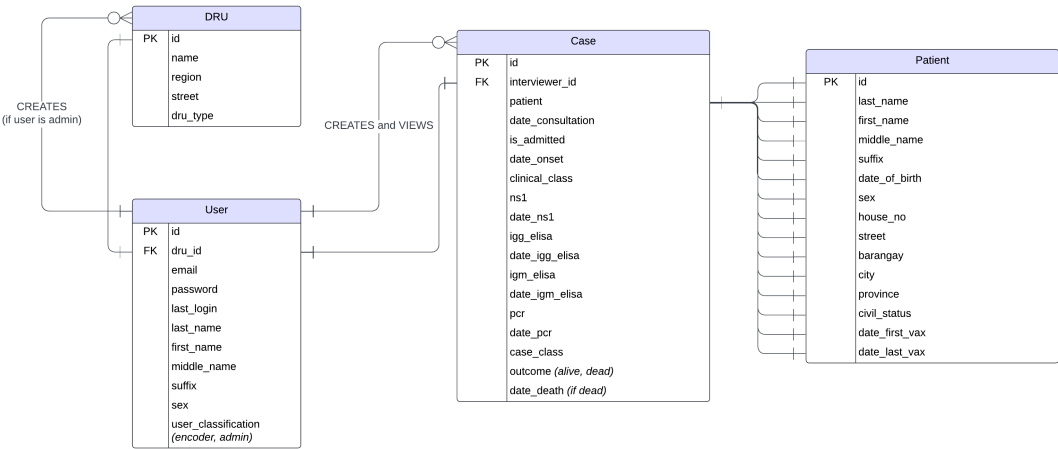


Figure 3.3: Entity-Relationship Database Schema Hybrid Diagram for DengueDash Database Structure

3.3.2 User Interface Requirements

Admin Interface

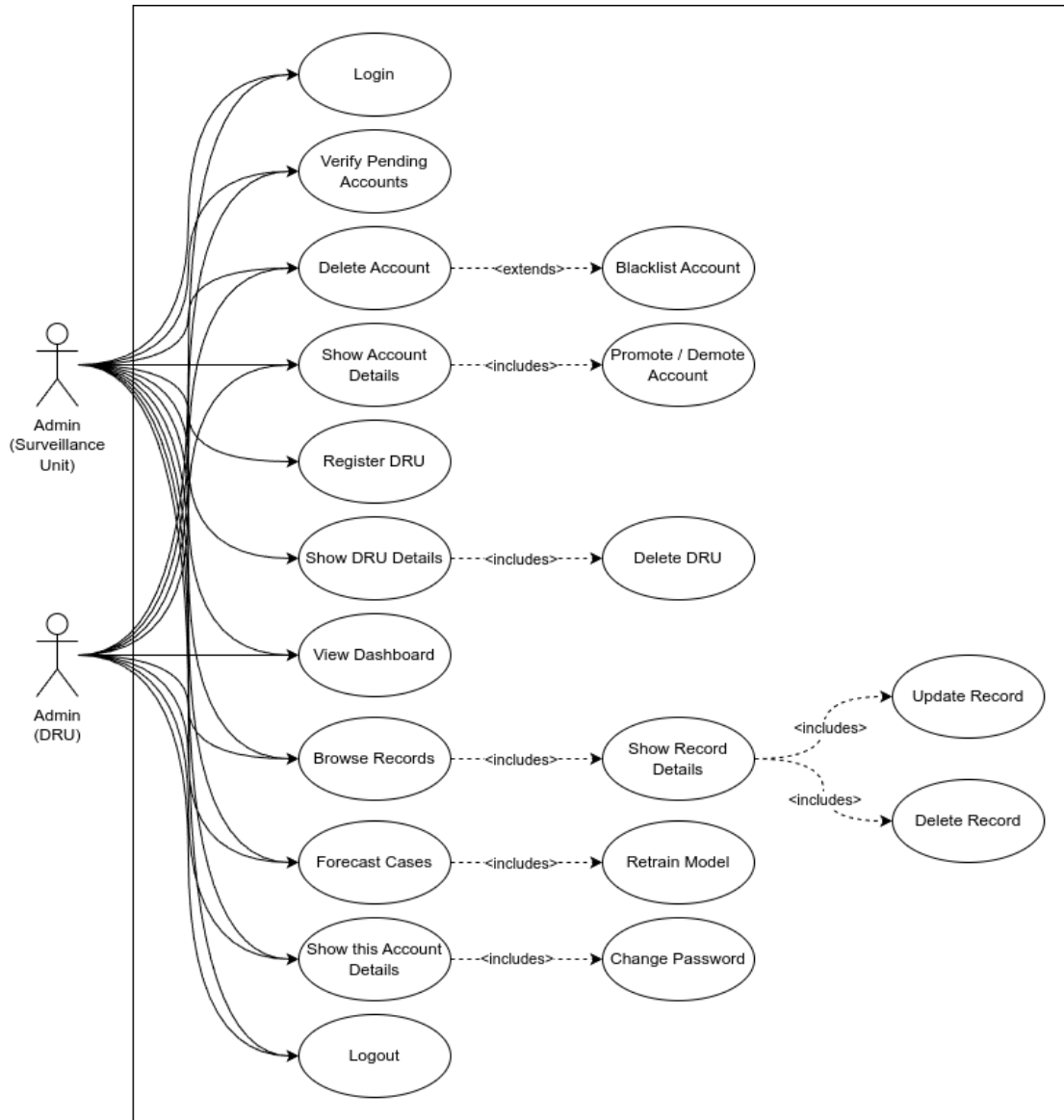


Figure 3.4: Use Case Diagram for Admins

Figure 3.4 shows the actions of an admin for a specific Disease Reporting Unit (DRU) and an admin for a specific Surveillance Unit can take in the application. Both of them include the management of accounts, browsing records, and forecasting and retraining all the consolidated data under their supervision. Most

719 importantly, these users must verify the encoders who register under their ju-
 720 risdiction before allowing their account to access the application in the name of
 721 safeguarding the integrity of the data. The only advantage of the latter type of ad-
 722 ministrator is that it has a one-step higher authorization as it manages the DRUs.
 723 In addition, only the authorized surveillance unit administrator can register and
 724 create a DRU to uphold transparency and accountability.

725 Encoder Interface

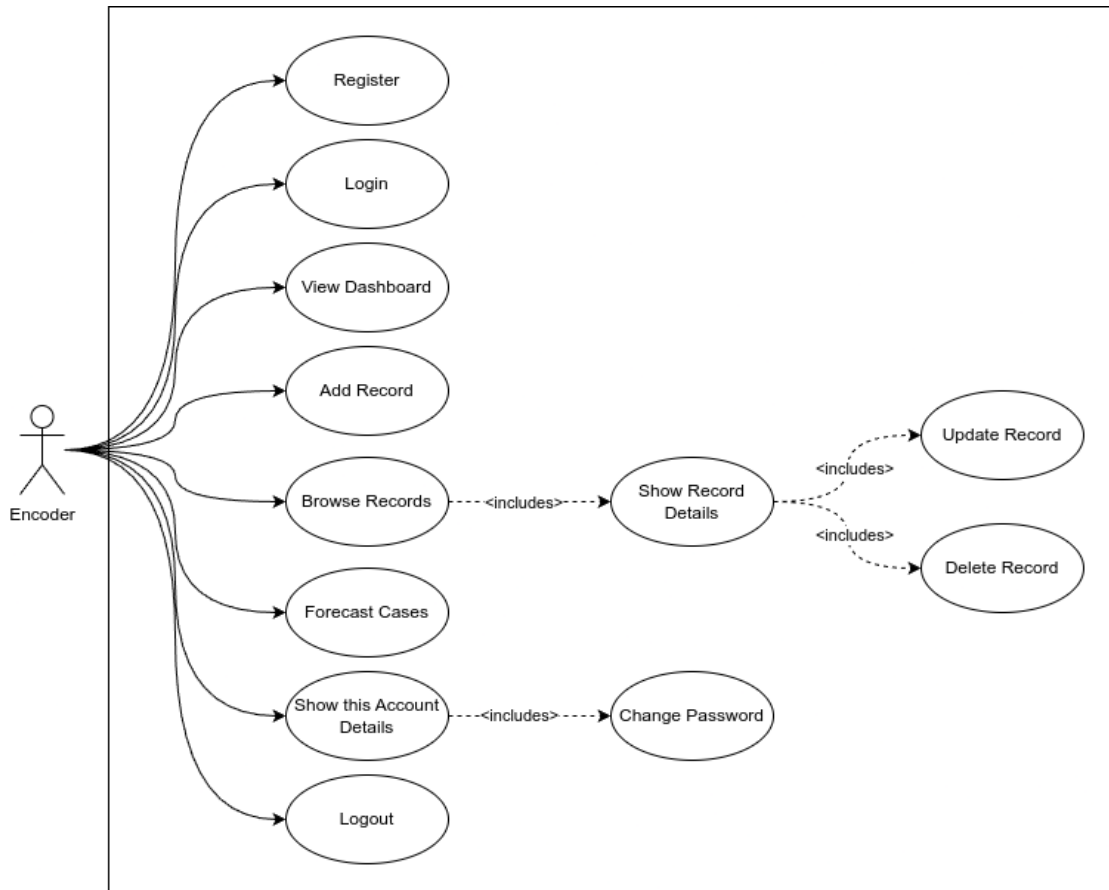


Figure 3.5: Use Case Diagram for Encoder

726 Figure 3.5, on the other hand, illustrates the use cases for the system's primary
 727 users. These users can register but must wait for further verification to access the
 728 application. Similar to the previous interfaces, encoders can browse and manage
 729 records, as well as forecast the consolidated cases under a specific surveillance or
 730 disease reporting unit, but they are not allowed to retrain the model. Lastly, they

731 are the only type of user that can file and create dengue cases by filling out a form
732 with the required details.

733 **3.3.3 Security and Validation Requirements**

734 **Password Encryption**

735 Storing passwords as plain text in the database is a disgrace and a mortal sin in
736 production. It is important to implement precautionary methods such as hashing
737 and salting, followed by encryption with a strong algorithm, to prevent bad actors
738 from using the accounts for malicious transactions. By default, Django generates
739 a unique random salt for each password and encrypts it with Password-Based Key
740 Derivation Function 2 (PBKDF2) with a SHA256 hash function. Utilizing these
741 techniques ensures that in the event of a data breach, cracking these passwords
742 would be time-consuming and useless for the attackers.

743 **Authentication**

744 DengueWatch utilizes JSON Web Tokens (JWT) to authenticate the user. Since
745 the mechanism operates in a stateless manner, tokens are served only after a
746 successful login, eliminating the need for the server to keep a record of the token,
747 which is vulnerable to session hijacking. In addition, these tokens are signed with
748 a secret key, ensuring they have not been tampered with.

749 **Data Validation**

750 Both the backend and frontend should validate the input from the user to preserve
751 data integrity. Thus, Zod is implemented in the latter to help catch invalid inputs
752 from the user. By doing this, the user can only send proper requests to the server
753 which streamlines the total workflow. On the other hand, Django has also a built-
754 in validator that checks the data type and ensures that the input matches the
755 expected format on the server side. These validation processes ensure that only
756 valid and properly formatted data is accepted, which reduces the risk of errors
757 and ensures consistency across the web application.

3.4 Calendar of Activities

A Gantt chart showing the schedule of the activities is included below. Each bullet represents approximately one week of activity.

Table 3.1: Timetable of Activities for 2024

Activities	Aug	Sept	Oct	Nov	Dec
Project Initiation and Team Formation	••				
Literature Review and Data Gathering	••	••••			
Data Cleaning and Feature Selection		••		•	•
Creating System Dashboard		••	••••	•	
Analysis and Interpretation of Results			•		•
Documentation	••	••••	••••	••••	••••

Table 3.2: Timetable of Activities for 2025

Activities	Jan	Feb	Mar	Apr	May
Create Admin Dashboard	•	•••			
Integrate the Best Model to the System	•	••••			
Extend Features to Accommodate a National Setting		•	••		
User Testing			••	•	
Documentation	••	••••	••••	••••	••••

Chapter 4

Results and Discussion/System Prototype

4.1 Data Gathering

The data for dengue case prediction was gathered from a variety of reliable sources, enabling a comprehensive dataset spanning from January 2011 to October 2024. This dataset includes 720 rows of data, each containing weekly records of dengue cases along with corresponding meteorological variables, such as rainfall, temperature, and humidity.

1. Dengue Case Data: The primary source of historical dengue cases came from the Humanitarian Data Exchange and the Western Visayas Center for Health Development (WVCHD). The dataset, accessed through Freedom of Information (FOI) requests, provided robust case numbers for the Western Visayas region. The systematic collection of these data points was essential for establishing a reliable baseline for model training and evaluation.
2. Weather Data: Weekly weather data was obtained by web scraping from Weather Underground, allowing access to rainfall, temperature, wind, and humidity levels that correlate with dengue prevalence.

```
data.head()
```

	Time	Rainfall	MaxTemperature	AverageTemperature	MinTemperature	Wind	Humidity	Cases
0	2011-01-03	9.938571	29.444400	25.888890	23.888900	11.39	86.242857	5
1	2011-01-10	8.587143	30.000000	26.705556	24.444444	7.32	88.028571	4
2	2011-01-17	5.338571	30.000000	26.616667	25.000000	7.55	84.028571	2
3	2011-01-24	5.410000	30.555556	26.483333	20.555556	10.67	80.971429	7
4	2011-01-31	2.914286	28.333333	25.283333	18.650000	11.01	74.885714	2

Figure 4.1: Snippet of the Combined Dataset

#	Column	Non-Null	Count	Dtype
---	-----	-----	-----	-----
0	Time	720	non-null	datetime64[ns]
1	Rainfall	720	non-null	float64
2	MaxTemperature	720	non-null	float64
3	AverageTemperature	720	non-null	float64
4	MinTemperature	720	non-null	float64
5	Wind	720	non-null	float64
6	Humidity	720	non-null	float64
7	Cases	720	non-null	int64
dtypes: datetime64[ns](1), float64(6), int64(1)				
memory usage: 45.1 KB				

Figure 4.2: Data Contents

779 4.2 Exploratory Data Analysis

780 From the summary above, the dataset consists of 720 weekly records with 8
781 columns:

- 782 • **Time.** Weekly timestamps (e.g. "2011-w1")
- 783 • **Rainfall.** Weekly average rainfall (mm)
- 784 • **MaxTemperature, AverageTemperature, MinTemperature.** Weekly
785 temperature data (C)
- 786 • **Wind.** Wind speed (m/s)
- 787 • **Humidity.** Weekly average humidity (%)
- 788 • **Cases.** Reported dengue cases

	Time	Rainfall	MaxTemperature	AverageTemperature	MinTemperature	Wind	Humidity	Cases
count	720	720.000000	720.000000	720.000000	720.000000	720.000000	720.000000	720.000000
mean	2017-12-02 11:22:00	13.957499	32.191142	28.110319	25.038472	6.172417	81.609442	23.744444
min	2011-01-03 00:00:00	0.000000	-14.600000	24.494444	12.222222	1.910000	73.185714	0.000000
25%	2014-06-21 06:00:00	1.270000	31.666667	27.504167	25.000000	4.117500	79.885713	5.000000
50%	2017-12-07 12:00:00	4.318000	32.222222	28.161111	25.000000	5.725000	81.771429	12.000000
75%	2021-05-11 18:00:00	10.414000	32.777778	28.751389	25.555556	7.860000	83.503571	26.000000
max	2024-10-28 00:00:00	445.008000	58.333333	30.916667	32.222222	19.200000	89.571429	319.000000
std	NaN	35.448846	2.616379	0.999800	1.291659	2.446703	2.831674	37.144813

Figure 4.3: Dataset Statistics

From the statistics in figure 4.3, the number of cases ranges from 0 to 319. The average number of dengue cases per week is 23.74, with a median of 12 cases and a standard deviation of 37.14. The distribution is highly skewed, with some weeks experiencing significant number of cases (up to 319 cases). Rainfall shows a wide variation (0 to 445mm), while temperature remains relatively stable, with an average of 28.1 degree celsius. Humidity levels ranges from 73% to 89% with a mean of 81.6%.

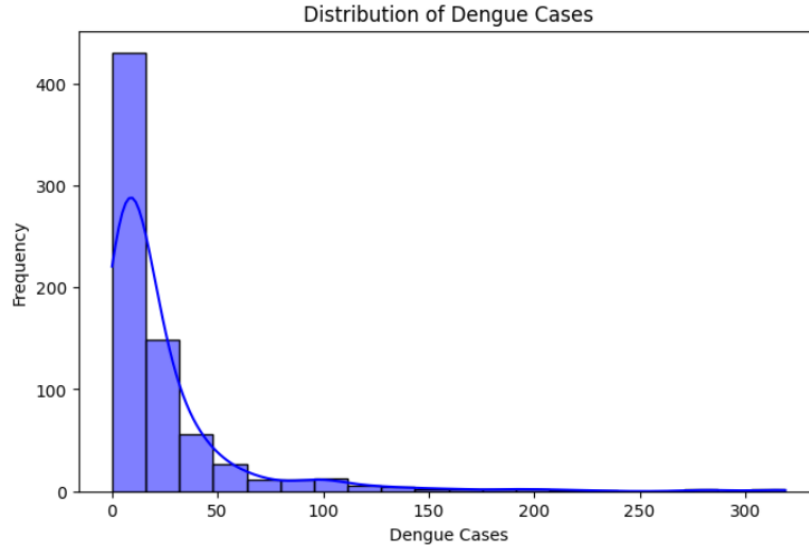


Figure 4.4: Distribution of Dengue Cases

In figure 4.4, a histogram of dengue cases shows a right-skewed distribution, indicating that most weeks experience low case counts, while a few weeks record outbreaks. To further analyze the distribution, dengue cases were categorized into different intervals (Figure 4.5): 0-5 cases, 6-15 cases, 16-30 cases, 31-100 cases and 101+

801 cases. The majority of weeks falls within the 0-5 cases and 6-15 cases categories,
 802 indicating that most weeks have low dengue cases. Meanwhile, weeks with 101+
 cases are rare, suggesting that extreme outbreaks are not frequent.

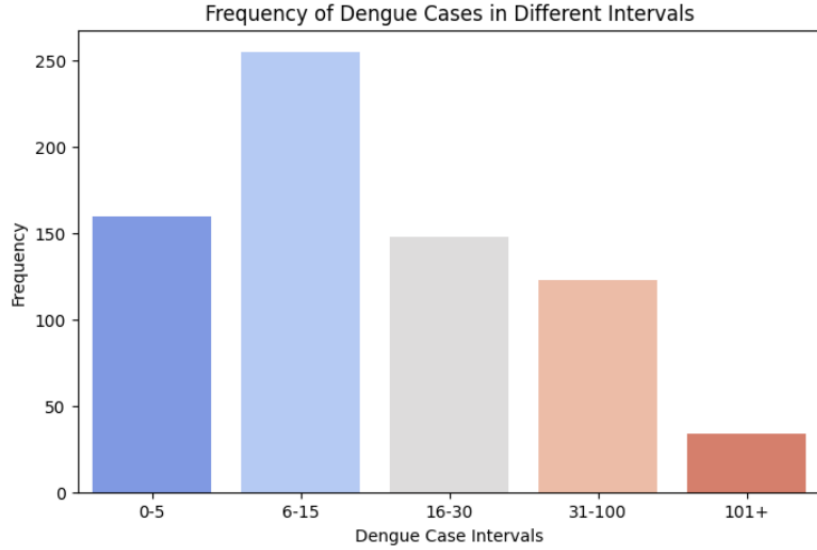


Figure 4.5: Frequency of Dengue Cases in Different Intervals

803

804 Figure 4.6 illustrates the trend of weekly dengue cases over time. The data
 805 reveals periodic spikes in the number of cases, suggesting a seasonal pattern in
 806 dengue cases. Notably, peak cases are observed during certain periods approx-
 807 imately 3 years, potentially aligning with specific climatic conditions such as
 808 increased rainfall or temperature changes. This underscores the importance of
 809 incorporating climate variables into the forecasting model.

810 Figure 4.7 shows the ranking of correlation coefficients between dengue cases
 811 and selected features, including rainfall, humidity, maximum temperature, aver-
 812 age temperature, minimum temperature, and wind speed. Among these, rainfall
 813 exhibits the highest positive correlation with dengue cases (correlation coefficient
 814 0.13), indicating that increased rainfall may contribute to higher cases counts.
 815 This aligns with existing studies suggesting that stagnant water from heavy rain-
 816 fall creates breeding grounds for mosquitos. It is followed by humidity (0.10),
 817 suggesting that higher humidity levels may enhance mosquito reproduction, lead-
 818 ing to more dengue cases. Temperature has a weak to moderate positive corre-
 819 lation with dengue cases, with maximum temperature (0.09) showing a stronger
 820 relationship than average and minimum temperature.

821 Figure 4.8 shows the ranking of correlation coefficients between dengue cases
 822 and selected features, with the addition of lagged effects. The analysis reveals no

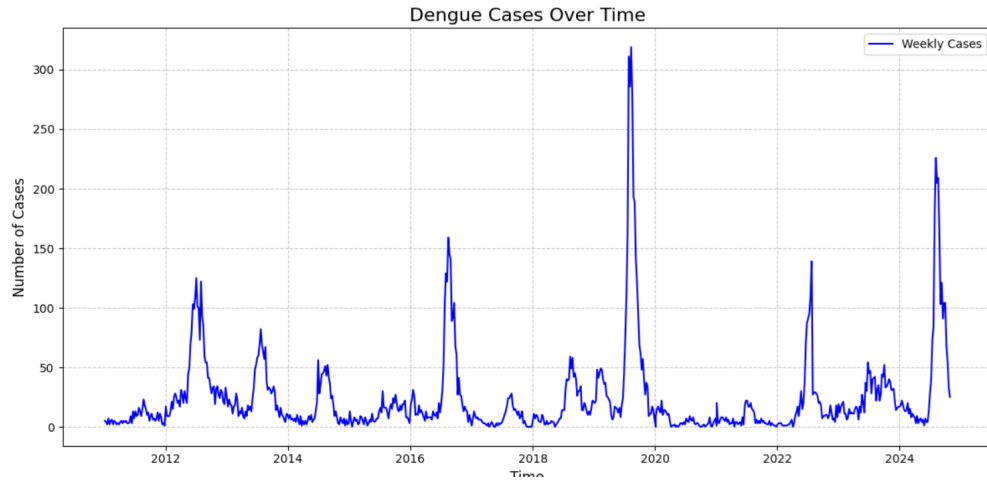


Figure 4.6: Trend of Dengue Cases

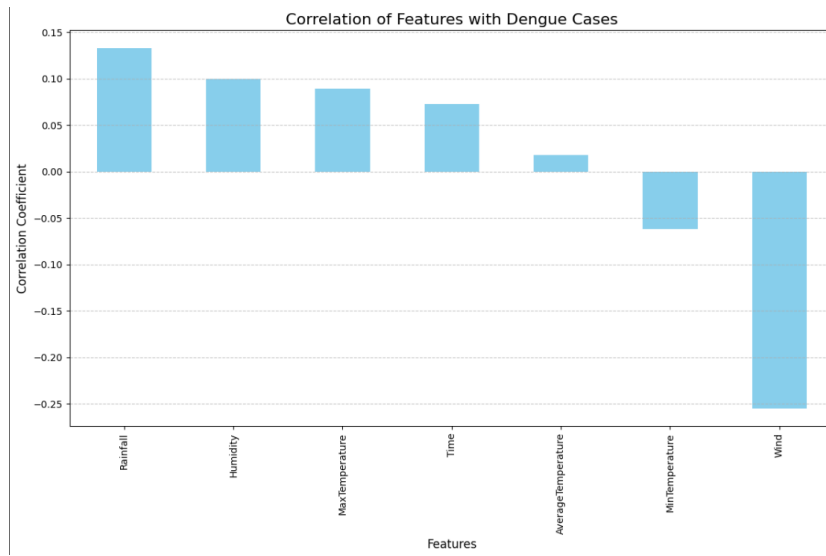


Figure 4.7: Ranking of Correlations

improvement in correlation when lagged variables are compared to direct observations. This suggests that the observed values of rainfall, humidity, and maximum temperature remain the most significant predictors for dengue case forecasting. Overall, the exploratory data analysis highlights the significance of rainfall, humidity, and max temperature variables in dengue case forecasting.

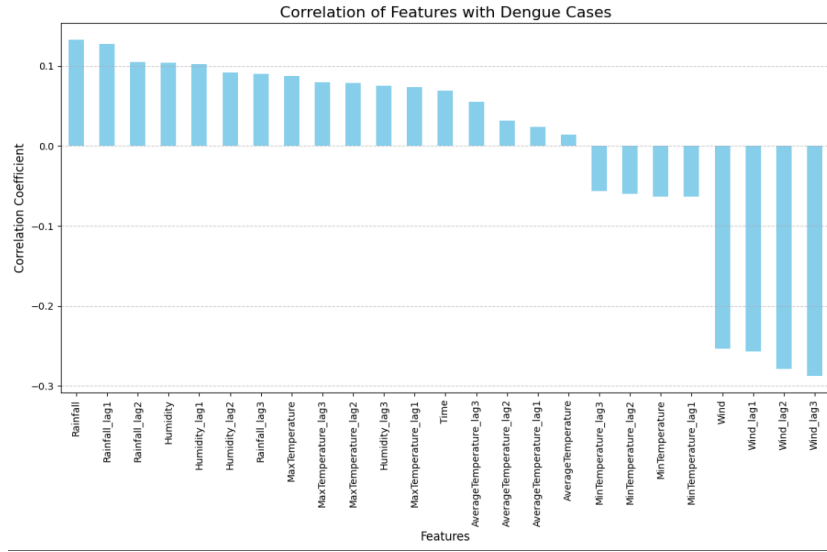


Figure 4.8: Ranking of Correlations (with lagged effects)

4.3 Outbreak Detection

To identify outbreaks, we calculated the outbreak threshold value using the historical mean as the endemic channel. The threshold is determined using the formula:

$$\text{Outbreak Threshold Value} = \mu + 2\sigma \quad (4.1)$$

$$= 23.744444 + 2(37.144813) \quad (4.2)$$

$$= 23.744444 + 74.289626 \quad (4.3)$$

$$= 98.03407 \quad (4.4)$$

where μ is the historical mean and σ is the standard deviation.

This result indicates that dengue cases exceeding 98 in Iloilo City can be considered an outbreak. However, it is important to note that this threshold serves only as a baseline. Additional parameters, such as the number of hospital beds available in the city, must be considered to compute a more effective threshold and develop an appropriate response strategy.

837 4.4 Model Training Results

838 The models were evaluated using three metrics: MSE, RMSE, and MAE. The
 839 table below provides a summary and comparative analysis of each model's results
 840 across these metrics, offering insights into the strengths and limitations of each
 841 forecasting technique for dengue case prediction in Iloilo City. The lower values
 842 of the three metrics indicate better forecasting performance. Table 4.1 shows that
 843 the models performed differently on testing data. LSTM outperformed the other
 844 models with the lowest RMSE, MSE, and MAE while the other three models had
 845 relatively higher values for the three metrics.

Method	LSTM	Seasonal ARIMA	ARIMA	Kalman Filter	KF-LSTM
Testing MSE	285.54	1261.20	1521.48	1474.82	785.35
Testing RMSE	16.90	34.45	39.00	38.40	25.56
Testing MAE	9.45	18.73	25.80	22.33	14.55
Best Parameters	Window Size: 5 Learning Rate: 0.01 Units: 64	(2,0,2)(0,1,1)	(1,2,2)	Observation Covariance: 10.0 Transition Covariance: $0.1 \times \text{Identity}$	Same as LSTM

Table 4.1: Comparison of different models for dengue prediction

846 4.4.1 LSTM Model

847 The LSTM model was tuned for the following parameters: learning rate and units.
 848 The hyperparameter tuning was conducted for each window size, finding the best
 849 parameters for each window size. Further evaluating which window size is most
 850 suitable for the prediction model, Table 4.2 shows the evaluation metrics for each
 window size used in the LSTM model training.

Window Size	MSE	RMSE	MAE	R ²
5	285.54	16.90	9.45	0.83
10	334.63	18.29	9.85	0.80
20	294.85	17.17	9.35	0.83

Table 4.2: Comparison of Window Sizes

851

852 The results indicate that a window size of 5 weeks provides the most accurate
 853 predictions, as evidenced by the lowest MSE and RMSE values. Furthermore, the
 854 R² score of 0.83 indicates that 83% of the variability in the target variable (cases)
 855 is explained by the independent variables (the inputs) in the model, making it a
 856 reliable configuration overall.

857 Figure 4.9 illustrates the model's performance in predicting dengue cases for
 858 each fold using a window size of 5. As shown in the plot, the training set progres-

859 sively increases with each fold, mimicking a real-world scenario where more data
 860 becomes available over time for dengue prediction. Figure 4.10 demonstrates that
 861 the predicted cases closely follow the trend of the actual cases, indicating that the
 862 LSTM model successfully captures the underlying patterns in the data. It is also
 863 evident that as the fold number increases and the training set grows, the accuracy
 864 of the predictions on the test set improves. Despite the test data being unseen,
 865 the model exhibits a strong ability to generalize, suggesting it effectively leverages
 866 past observations to predict future trends.

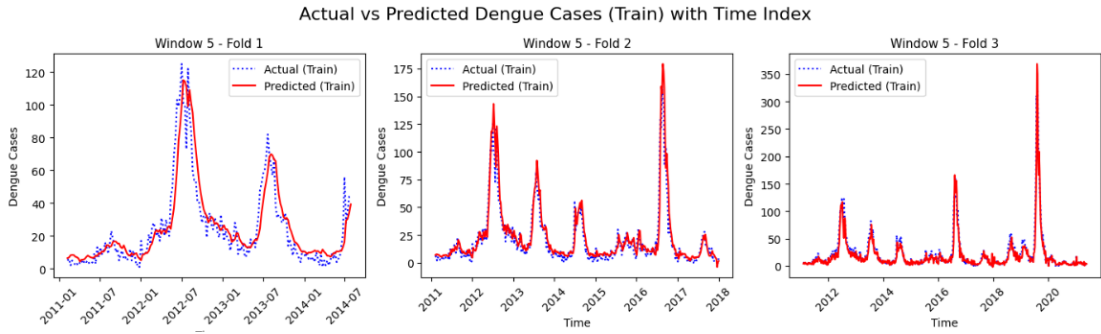


Figure 4.9: Training Folds - Window Size 5

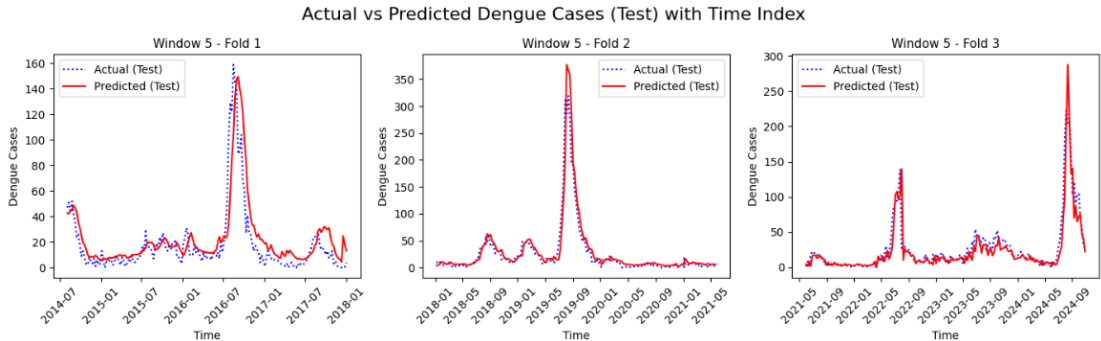


Figure 4.10: Testing Folds - Window Size 5

867 4.4.2 ARIMA Model

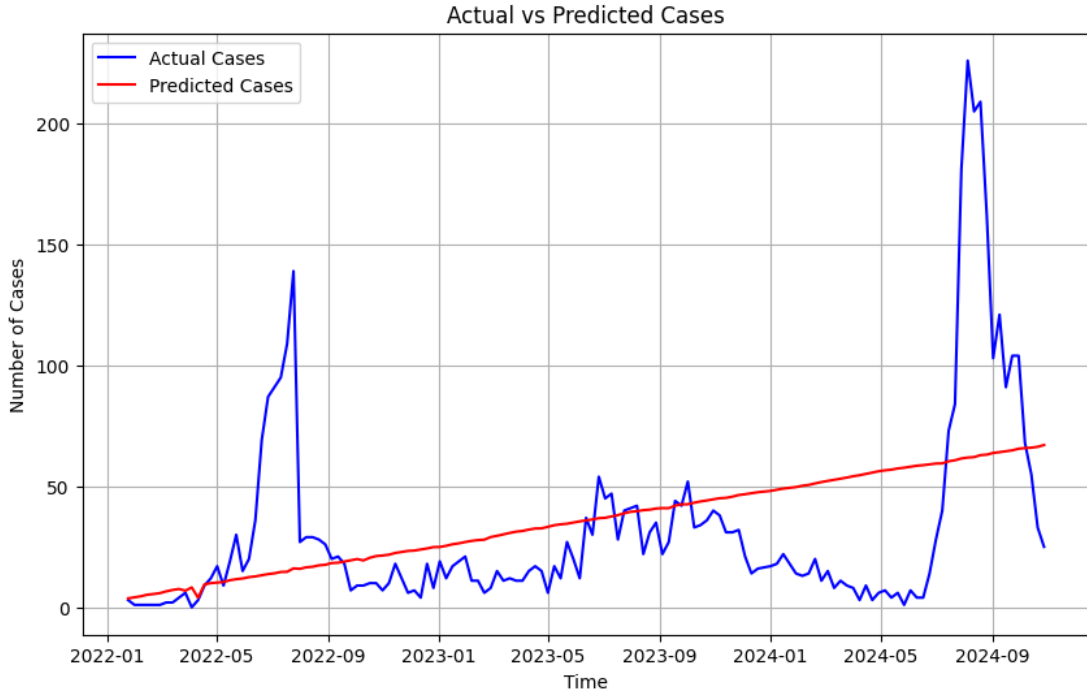


Figure 4.11: ARIMA Prediction Results for Test Set

868 The ARIMA model was developed to capture non-seasonal trends in the data.
 869 To determine the best model configuration, grid search was used to explore vari-
 870 ous combinations of ARIMA parameters, ultimately selecting **ARIMA(1, 2, 2)**.
 871 The model was iteratively refined over **400 iterations** to ensure convergence to
 872 an optimal solution. Figure 4.11 illustrates the comparison between actual and
 873 predicted dengue cases in the test set. As shown in the plot, the ARIMA model
 874 struggled to capture the non-linear characteristics and abrupt spikes in the data.
 875 Consequently, it failed to accurately reflect the fluctuations and outbreak patterns
 876 seen in the actual case counts.

877 The model's performance was assessed using regression metrics to evaluate its
 878 forecasting capability. The ARIMA model yielded the following error metrics:

- 879 • **MSE (Mean Squared Error):** 1521.48
- 880 • **RMSE (Root Mean Squared Error):** 39.01
- 881 • **MAE (Mean Absolute Error):** 25.80

4.4.3 Seasonal ARIMA (SARIMA) Model

To address the limitations of the ARIMA model, a Seasonal ARIMA (SARIMA) model was developed to capture both non-seasonal and seasonal variations in the data.

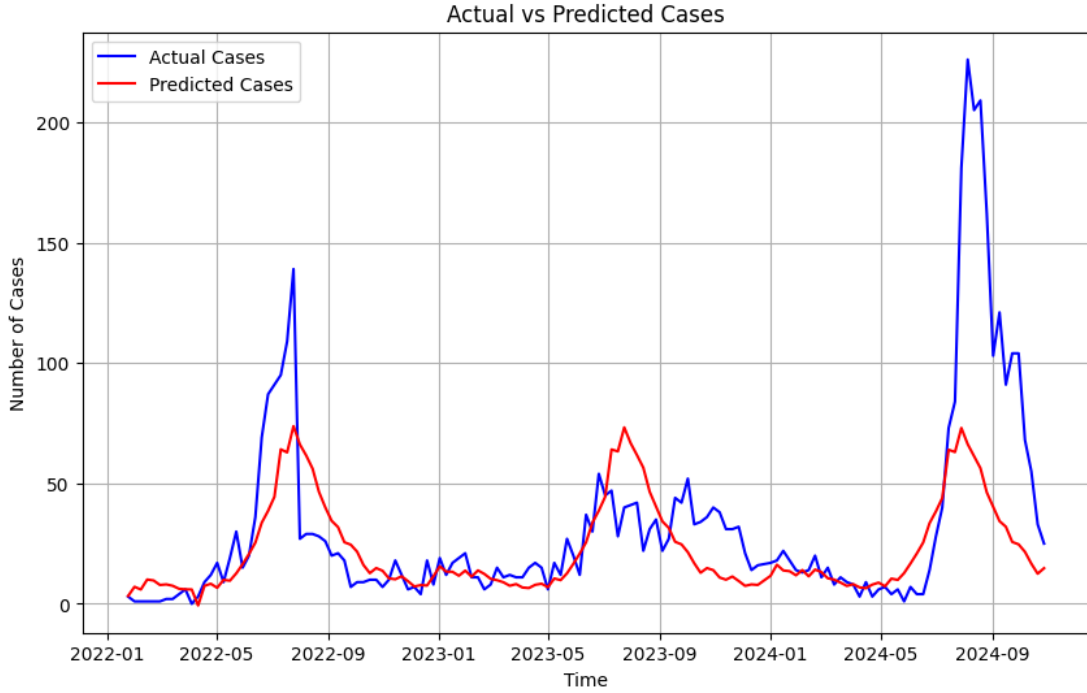


Figure 4.12: Seasonal ARIMA Prediction Results for Test Set

This model incorporates seasonal parameters, which were tuned using grid search to find the best configuration: **SARIMA(2, 0, 2)(0, 1, 1)[52]**. As with ARIMA, **400 iterations** were applied to ensure a robust fit. As shown in Figure 4.12, the SARIMA model demonstrates a notable improvement in performance. Unlike its non-seasonal counterpart, it effectively captures the general trend and aligns more closely with the peaks observed in the actual dengue cases, indicating its ability to model seasonal dynamics.

The model's performance was assessed using regression metrics to evaluate its forecasting capability. The SARIMA model yielded the following error metrics:

- **MSE:** 1109.69
- **RMSE:** 33.31

897

- **MAE: 18.09**

898 The lower error values, when compared to the ARIMA model, highlight the
 899 SARIMA model's superior capability in forecasting dengue cases. Its effective-
 900 ness in capturing seasonal patterns contributed to a more accurate representation
 901 of the actual cases.

902 After training the model, the SARIMA model was validated using the same
 903 Time Series Cross-Validation strategy employed in the LSTM model. Table 4.3
 904 presents the performance metrics for each fold, as well as the average metrics
 905 across all folds. The average RMSE and MAE values were close to those obtained
 906 during the initial training phase, indicating that the SARIMA model performed
 907 consistently across different time segments.

Fold	MSE	RMSE	MAE
1	659.68	25.68	16.00
2	2127.49	46.12	21.30
3	996.43	31.56	18.89
Average	1261.20	34.45	18.73

Table 4.3: Comparison of SARIMA performance for each fold

908 4.4.4 Kalman Filter Model

909 Figure 4.13 shows the comparison between the actual dengue cases and the pre-
 910 dicted values on the test set. As illustrated in the plot, the Kalman Filter model
 911 demonstrates a moderate ability to follow the general trend of the actual data.
 912 While it effectively captures some rising and falling patterns, it still struggles to
 913 accurately replicate the sharp peaks and extreme values found in the real case
 914 counts. This limitation is particularly noticeable during the large spikes in 2022
 915 and 2024. The model's performance was evaluated using standard regression met-
 916 rics. The results are as follows:

$$\text{MSE} = 1474.82, \quad \text{RMSE} = 38.40, \quad \text{MAE} = 22.34$$

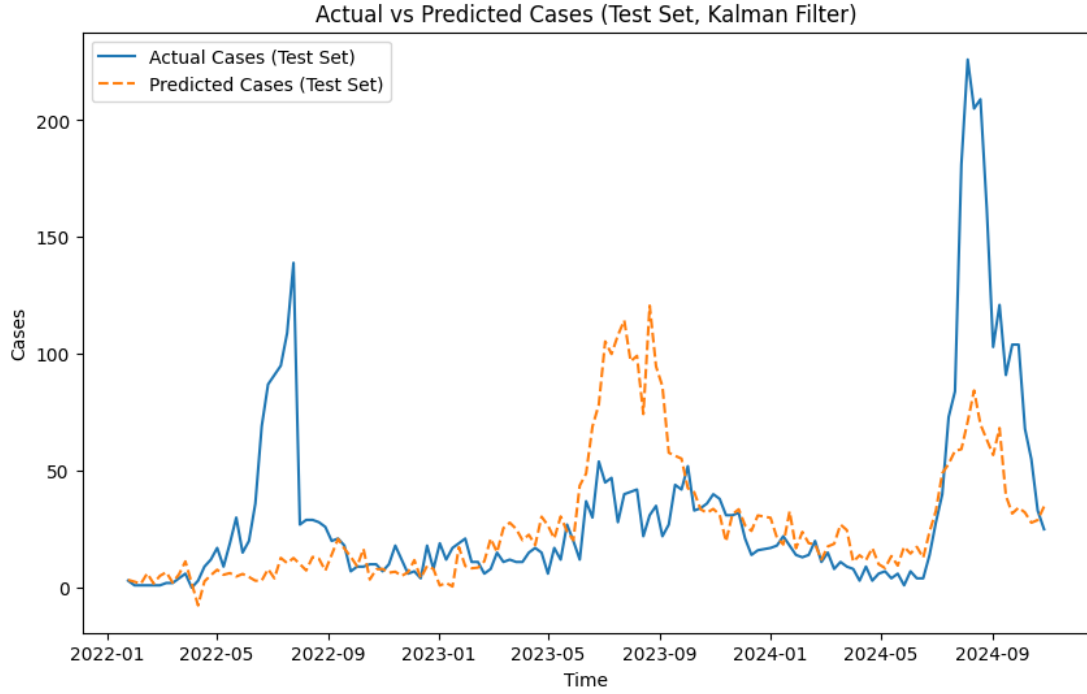


Figure 4.13: Kalman Filter Prediction Results for Test Set

917 The Kalman Filter was then combined with the LSTM model in order to see
 918 improvements in its predictions. Table 4.4 shows the metrics across three folds
 919 using the same Time Series Cross Validation Strategy employed in the previous
 920 models to see how it performed on different time segments.

Fold	MSE	RMSE	MAE
1	113.59	10.66	6.42
2	752.51	27.43	12.11
3	1489.95	38.60	25.13
Average	785.35	25.56	14.55

Table 4.4: Comparison of KF-LSTM performance for each fold

921 As can be seen in the table above, the performance of the hybrid model demon-
 922 strated improvements in all metrics as compared to just using the Kalman Filter
 923 alone.

4.5 Model Simulation

To evaluate the LSTM model's real-world forecasting ability, a simulation was conducted to predict dengue cases for the year 2025. The model was trained exclusively on data from 2011 to 2024, using both dengue cases and weather variables. Importantly, the actual dengue case values for 2025 were never included during training. Instead, only the weather variables collected for 2025 were input into the model to generate predictions for that year. After prediction, the forecasted dengue cases for 2025 were compared against the true observed cases to assess the model's accuracy. Figure 4.14 shows that the predicted values closely follow the trend, although it may overestimate the dengue cases in some weeks.

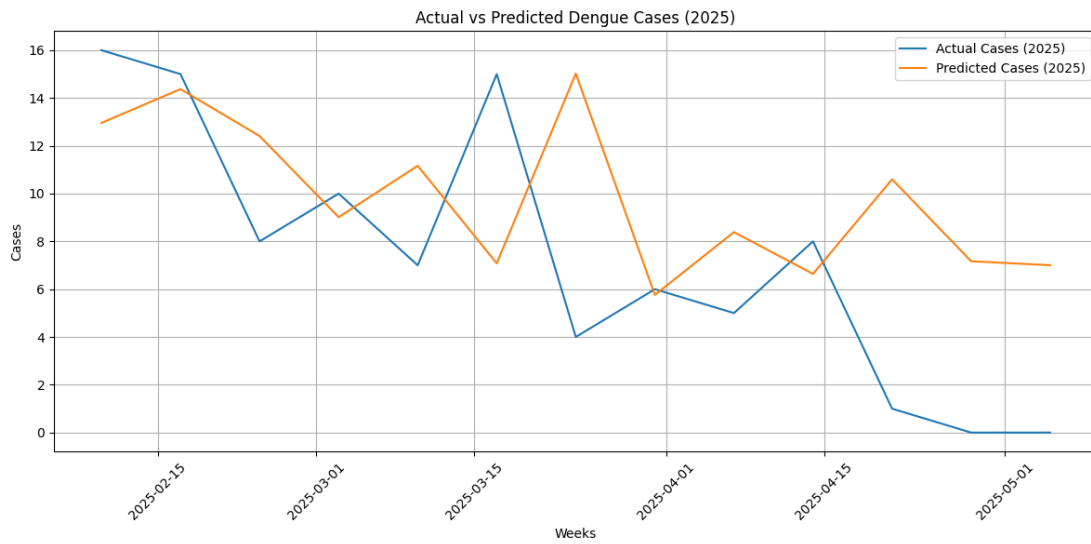


Figure 4.14: Predicted vs Actual Dengue Cases 2025

4.6 System Prototype

4.6.1 Home Page

The Home Page is intended for all visitors of the web application. The Analytics Dashboard, which displays relevant statistics for dengue cases at a certain year and location, is the primary component highlighted, as seen in Figure 4.15. This component includes a combo chart that graphs the number of dengue cases and deaths per week in a specific year, a choropleth map that tracks the number of

941 dengue cases per location, and various bar charts that indicate the top locations
 942 affected by dengue.

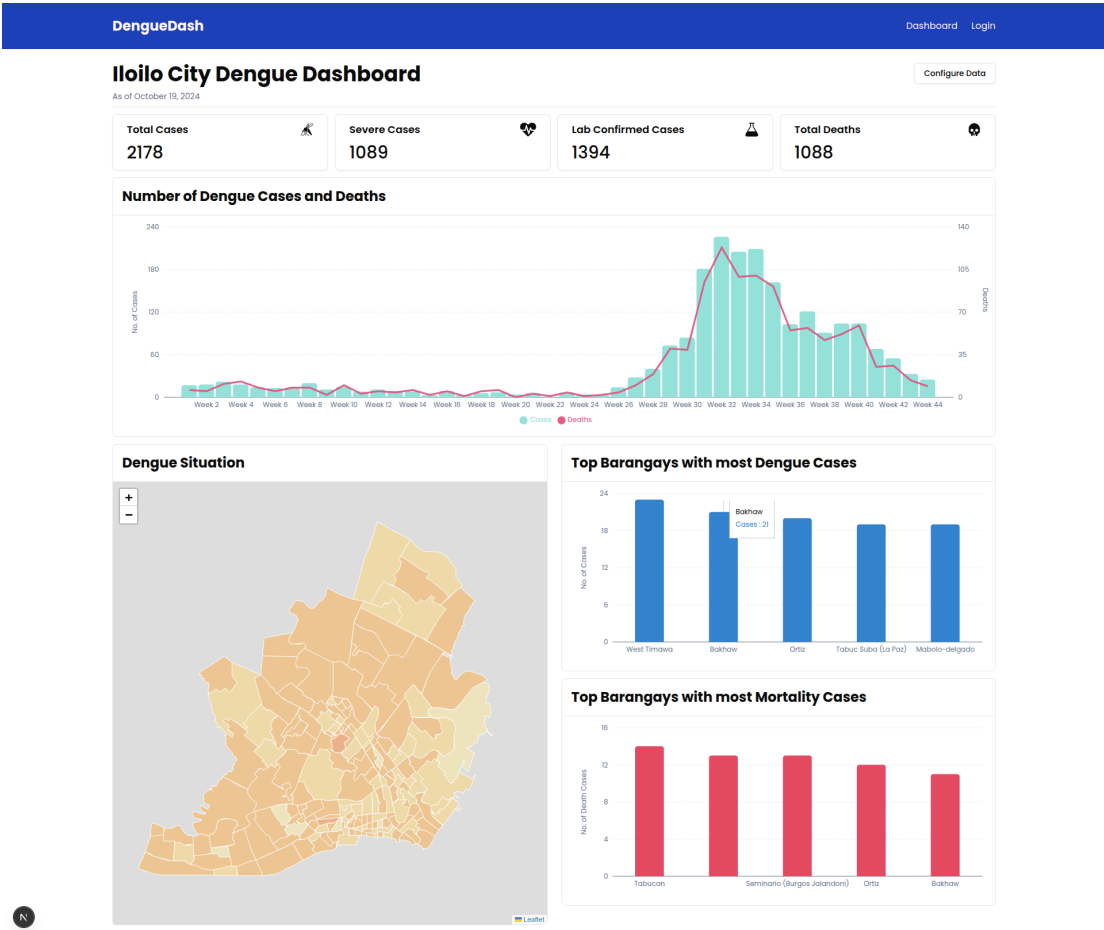


Figure 4.15: Home Page

943 4.6.2 User Registration, Login, and Authentication

944 The registration page, as shown in Figure 4.16, serves as a gateway to access the
 945 authenticated pages of the web application. Only prospected encoders can create
 946 an account since administrator accounts are only made by existing administra-
 947 tor accounts to protect the data's integrity in production. After registering, the
 948 "encoder account" cannot access the authorized pages yet as it needs to be veri-
 949 fied first by an administrator managing the unit the user entered. Once verified,
 950 the user can log in to the system through the page shown in Figure 4.17. Af-
 951 ter entering the correct credentials, which consist of an email and password, the

952 system uses HTTP-only cookies containing JSON Web Tokens (JWT) to prevent
953 vulnerability to Cross-site Scripting (XSS) attacks. It will then proceed to the
954 appropriate page the type of user belongs to.

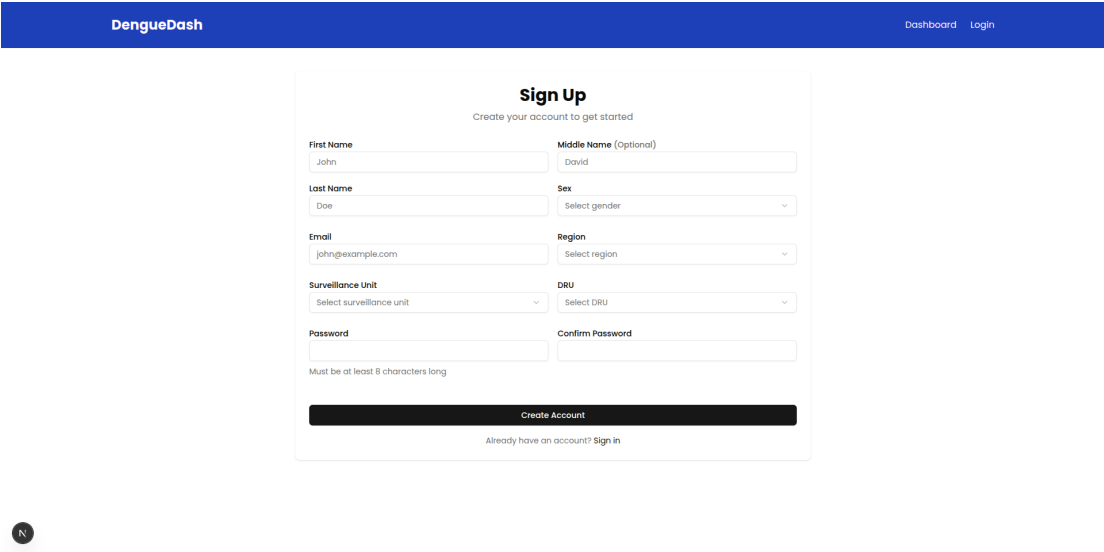


Figure 4.16: Sign Up Page

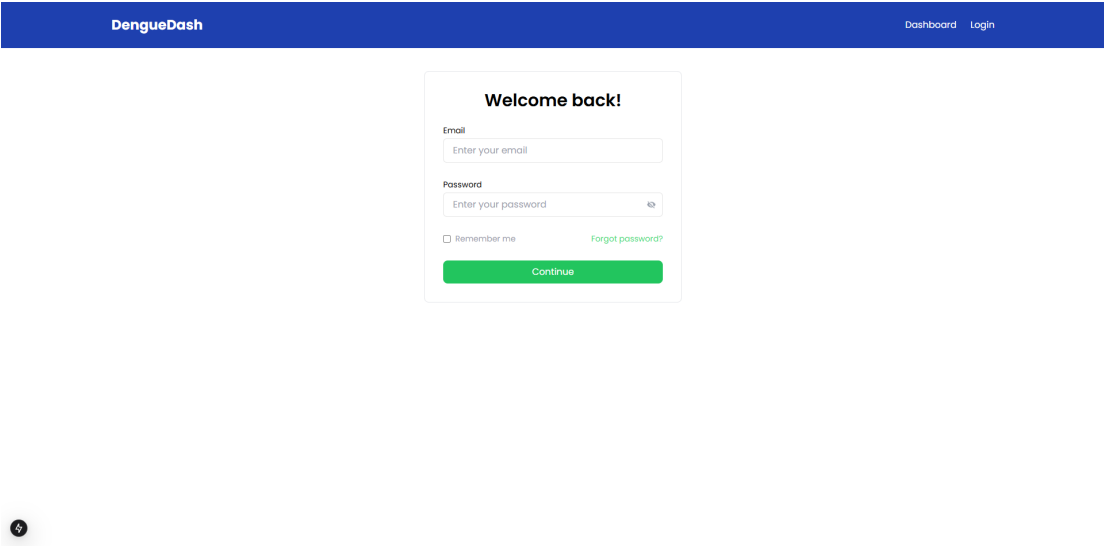


Figure 4.17: Login Page

4.6.3 Encoder Interface

Case Report Form

Figures 4.18 and 4.19 show the digitized counterpart of the form obtained from the Iloilo Provincial Epidemiology and Surveillance Unit. As the system aims to support expandability for future features, some fields were modified to accommodate more detailed input. It is worth noting that all of the included fields adhere to the latest Philippine Integrated Disease and Surveillance Response (PIDS) Dengue Forms, which the referenced form was based on. By doing this, if implemented on a national scale, the transition between targeted users will be easier. Moreover, the case form includes the patient's basic information, dengue vaccination status, consultation details, laboratory results, and the outcome. On the other hand, encoders can also create case records using a "bulk upload" feature that makes use of a formatted CSV file template. As shown in Figure 4.20, an encoder can download the template using the "Download Template" button, and insert multiple records inside the file, then upload it by clicking the "Click to upload" button. The web application automatically checks the file for data inconsistencies and validation.

The screenshot displays the 'Case Report Form' interface within the 'DengueDash' application. The left sidebar shows the navigation menu with options: Analytics, Forms (selected), Case Report Form, Data Tables, and Settings. The main content area is titled 'Case Report Form' and includes a 'Bulk Upload' button. The form is divided into two main sections: 'Personal Information' and 'Clinical Status'. The 'Personal Information' section contains fields for 'Personal Detail' (First Name, Middle Name, Last Name, Suffix, Sex, Civil Status, Date of Birth) and 'Address' (Region, Province, City, Barangay, Street, House No.). The 'Clinical Status' section contains fields for 'Vaccination' (Date of First Vaccination, Date of Last Vaccination). A 'Next' button is located at the bottom right of the form.

Figure 4.18: First Part of Case Report Form

DengueDash

Modules

Analytics

Forms

Case Report Form

Data Tables

Settings

Building Your Application > Data Fetching

Case Report Form

Bulk Upload

Personal Information

Clinical Status

Consultation

Date Admitted/Consulted/Seen

Pick a date

Is Admitted?

Select

Date Onset of illness

Pick a date

Clinical Classification

Select

Laboratory Results

NS1

Pending Result

Date done (NS1)

Pick a date

IgG ELISA

Pending Result

Date done (IgG ELISA)

Pick a date

IgM ELISA

Pending Result

Date done (IgM ELISA)

Pick a date

PCR

Pending Result

Date done (PCR)

Pick a date

Outcome

Case Classification

Select

Outcome

Select

Date of Death

Pick a date

Previous

Submit

Elizabeth Thomas Ro...

zewis@example.com

Figure 4.19: Second Part of Case Report Form

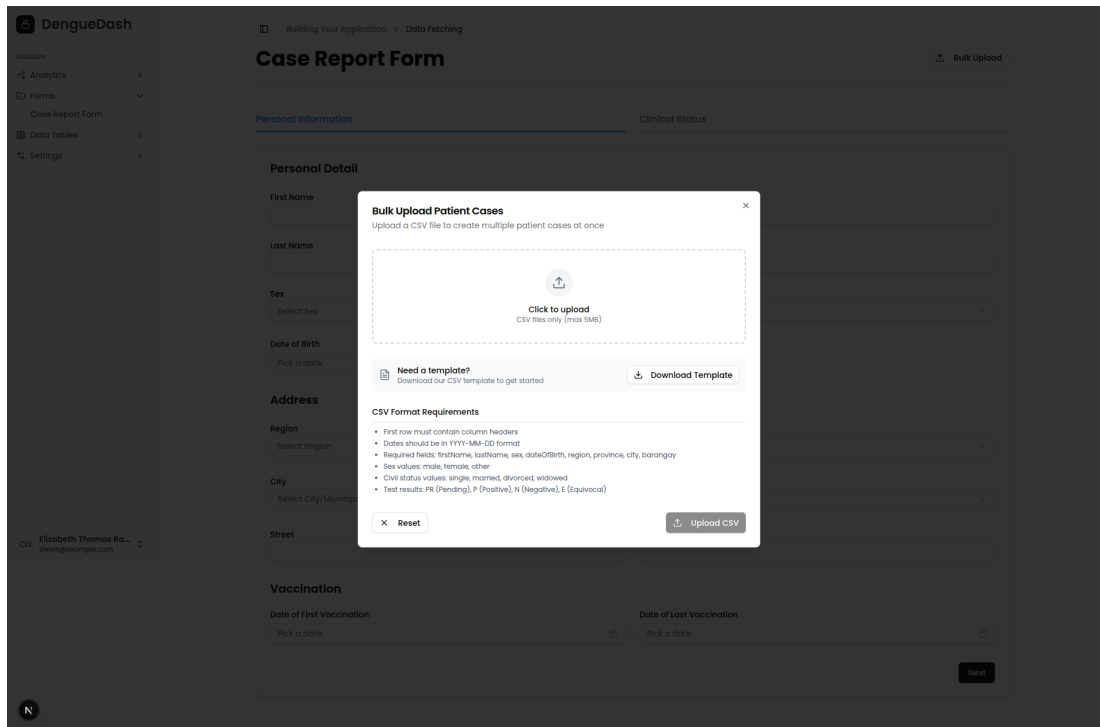


Figure 4.20: Bulk Upload of Cases using CSV

972 Browsing, Update, and Deletion of Records

973 Once the data generated from the case report form or the bulk upload is vali-
 974 dated, it will be assigned as a new case and can be accessed through the Dengue
 975 Reports page, as shown in Figure 4.21. The said page displays basic information
 976 about the patient related to a specific case, including their name, address, date
 977 of consultation, and clinical and case classifications. It is also worth noting that
 978 it only shows cases the user is permitted to view. For example, in a local Disease
 979 Reporting Unit (DRU) setting, the user can only access records that belong to
 980 the same DRU. On the other hand, in a consolidated surveillance unit such as a
 981 regional, provincial, or city quarter, its users can view all the records from all the
 982 DRUs that report to them. Moving forward, Figure 4.22 shows the detailed case
 983 report of the patient on a particular consultation date.

DengueDash

Modules

Accounts

>

DRU

>

Analytics

>

Data Tables

>

Dengue Reports

>

Settings

>

Ilolo City Epedemiol...

ilolocruz@gmail.com

Dengue Reports

Case ID	Name	Barangay	City	Date Consulted	Clinical Classification	Case Classification	Action
25017095	Hodges, Destiny Michelle	Pala Benedicto Rizal (Manduriao)	ILOILO CITY (Capital)	2024-11-03	Severe dengue	Probable	Open
25017077	Cuevas, Robert Rebecca	Democracia	ILOILO CITY (Capital)	2024-11-03	With warning signs	Confirmed	Open
25017090	Middleton, Joseph Michael	Tanza-Esperanza	ILOILO CITY (Capital)	2024-11-01	With warning signs	Suspect	Open
25017089	Medina, Michael Paige	Tacas	ILOILO CITY (Capital)	2024-11-01	With warning signs	Probable	Open
25017081	Love, Paula Kimberly	Magsaysay	ILOILO CITY (Capital)	2024-11-01	With warning signs	Suspect	Open
25017073	Smith, Anna Andrea	Desamparados	ILOILO CITY (Capital)	2024-11-01	Severe dengue	Confirmed	Open
25017094	Morrison, Michael Sarah	El 98 Castilla (Claudio Lopez)	ILOILO CITY (Capital)	2024-10-31	Severe dengue	Probable	Open
25017093	Barnes, Charles Robert	Rima-Rizal	ILOILO CITY (Capital)	2024-10-31	With warning signs	Suspect	Open

< Previous

12...

2137

Next >

Figure 4.21: Dengue Reports

Modules

- Accounts >
- DRU >
- Analytics >
- Data Tables >
- Dengue Reports >
- Settings >

Ilalo City Epidemiol...
ilalocase@gmail.com

Building Your Application > Data Fetching

Personal Information

Full Name
Medina, Michael Paige

Sex
Male

Full Address
995 Monique Spur, Tacas, ILOILO CITY (Capital), Ilalo

Date of Birth
October 11, 1935

Civil Status
Widowed

Vaccination Status

First Dose
April 26, 2023

Last Dose
May 31, 2020

Case Record #25017089

Date of Consultation
November 1, 2024

Date Onset of Illness
October 23, 2024

Patient Admitted?
No

Clinical Classification
With warning signs

Laboratory Results

NSI Negative	Date Done October 27, 2024
IgG Elisa Equivocal	Date Done October 30, 2024
IgM Elisa Pending Result	Date Done N/A
PCR Pending Result	Date Done N/A

Outcome

Case Classification
Probable

Date of Death
October 31, 2024

Outcome
Dead

Interviewer

Interviewer
Daniels, Lisa Long

DRU
Molo District Health Center

Update Case

Delete Case

Figure 4.22: Detailed Case Report

984 To update the case, the user can click the "Update Case" button, where a
 985 dialog will appear, and the updateable fields will be shown. It is worth noting
 986 that in this case, only fields under Laboratory Results and Outcome are included
 987 since they are the only ones that are time-based, where the result may change in
 988 the future. After updating, a prompt will show confirming the action of the user.
 989 Moving forward, to delete a case record, the user must click the "Delete Case"
 990 button, and a prompt verifying the action will appear. After confirming, the case
 991 will be deleted permanently.

DengueDash

Building Your Application

Data Fetching

Modules

Accounts

DRU

Analytics

Data Tables

Dengue Reports

Settings

Personal Information

Full Name

Medina, Michael Paige

Date of Birth

October 11, 1935

Sex

Male

Civil Status

Widowed

Full Address

995 Monique Spur, Tacos, ILILO CITY (Capital), Iloilo

Vaccination Status

First Dose

April 26, 2023

Case Record #

Date of Consultation

November 1, 2024

Date Onset of Illness

October 23, 2024

Laboratory Results

NSI

Negative

IgG Elisa

Equivocal

IgM Elisa

Pending Result

PCR

Pending Result

Outcome

Case Classification

Probable

Date of Death

October 31, 2024

Interviewer

Daniels, Lisa Long

Update Case #25017095

Laboratory Results

NSI

Pending Result

IgG Elisa

Equivocal

IgM Elisa

Equivocal

PCR

Equivocal

Outcome

Case Classification

Probable

Date Done

n/a

Date Done

November 7th, 2024

Date Done

November 7th, 2024

Date Done

November 5th, 2024

Outcome

Outcome

Alive

Cancel

Save Changes

Update Case

Delete Case

Date Done

N/A

Outcome

Dead

Figure 4.23: Update Report Dialog

47

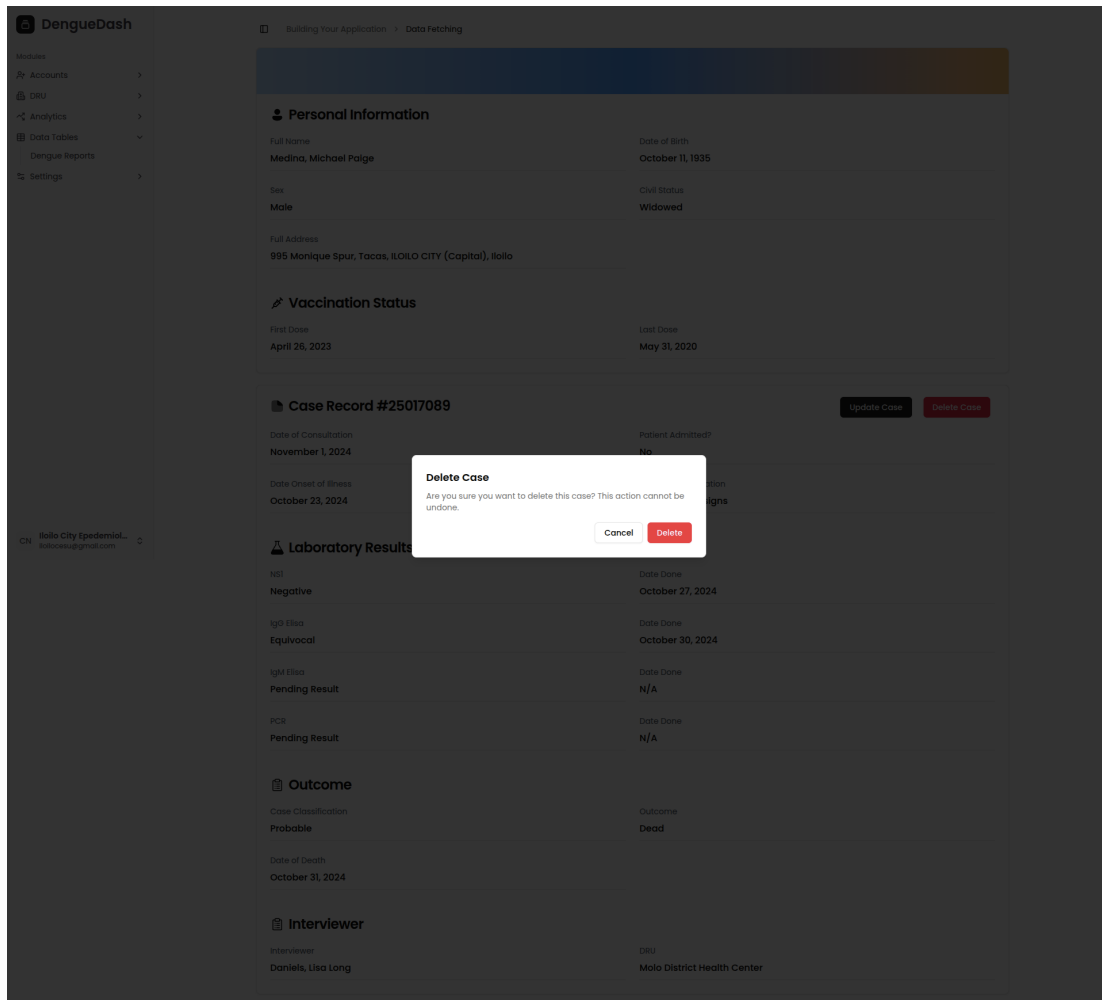


Figure 4.24: Delete Report Alert Dialog

992 Forecasting

993 The main highlight of the web application's feature is the Forecasting Page. This
 994 is where users can forecast dengue cases for the next following weeks. To predict,
 995 the application utilizes the exported LSTM model in a Keras format derived
 996 from training the consolidated data from the database. It requires the recent
 997 weekly dengue cases, weather variable data (temperature, humidity, and rainfall)
 998 based on the window size, and future weather data through OpenWeatherMap
 999 API. However, due to limitations imposed in the current plan subscribed in the
 1000 API, only the next 16 days of weather data can be fetched. As a result, the web
 1001 application can only make a two-week prediction. Moving forward, the Forecasting
 1002 page, as shown in Figure 4.25, introduces a user-friendly interface that shows the

1003 current cases for the week, and the predictions for the next two weeks with a range
1004 of 90 percent to 110 percent confidence interval that is presented in a simple but
1005 organized manner. There is also a line chart that shows the number of cases from
1006 the last 5 weeks plus the forecasted weekly cases. In addition, the current weather
1007 data for a specific week is also shown as well as the the forecasted weather data
1008 fetched from the said API. Lastly, locations where dengue cases have been reported
1009 for the current week are listed in the Location Risk Assessment component.

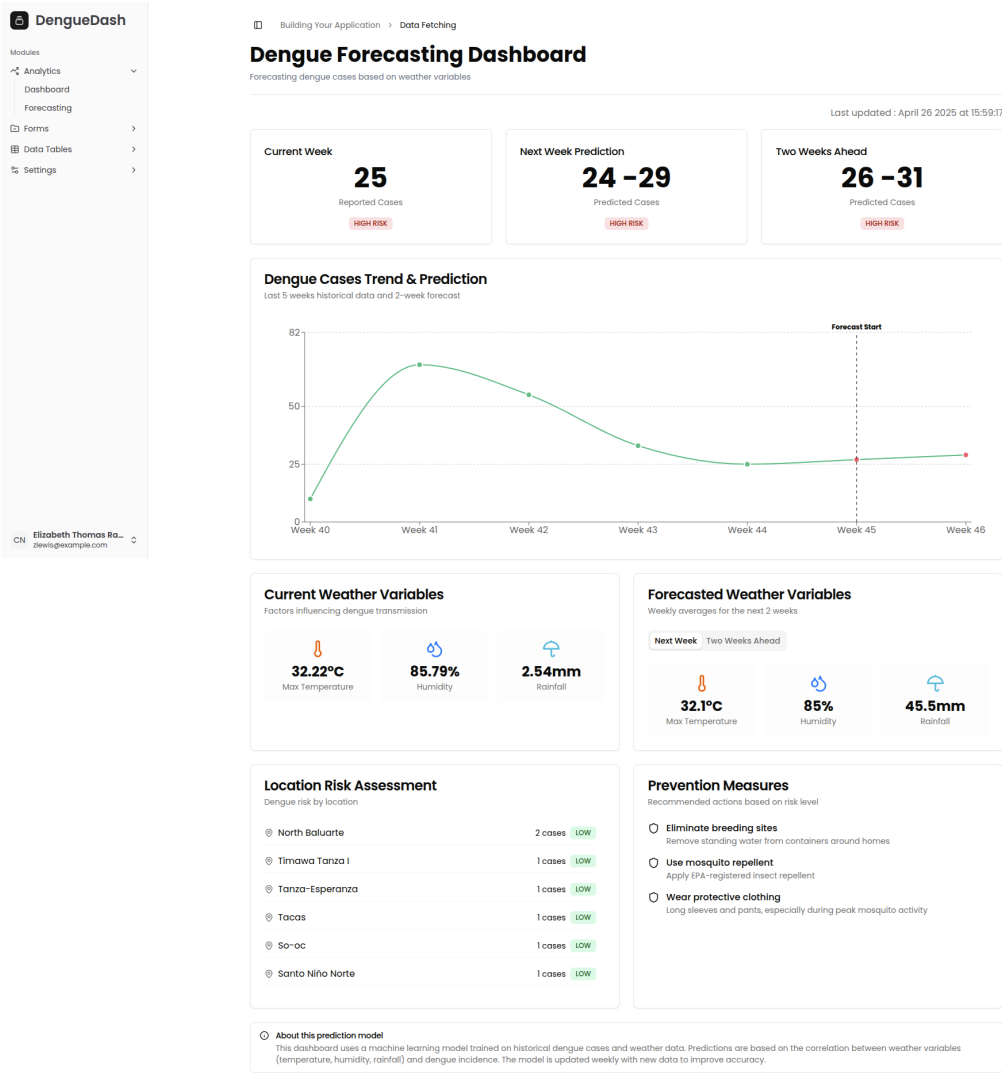


Figure 4.25: Forecasting Page

1010 **4.6.4 Admin Interface**

1011 **Retraining**

1012 With LSTM being the best-performing model among the models used in forecast-
1013 ing dengue cases, it is the model chosen to power the prediction and retraining
1014 of the consolidated data within the web application. Since the retraining process
1015 consumes a lot of processing power and requires a more advanced understanding
1016 of how it works, it was decided that the said feature should only be available
1017 to admin users. Furthermore, the retraining component in the Forecasting page
1018 includes three additional components that include the configuration of LSTM pa-
1019 rameters (Figure 4.26), the actual retraining of the consolidated data from the
1020 database (Figure 4.27), and the results of the retraining that shows the current
1021 and previous model metrics depending on the parameters entered (Figure 4.28).
1022 It is also worth noting that when trained, the model used a seeded number to
1023 promote reproducibility.

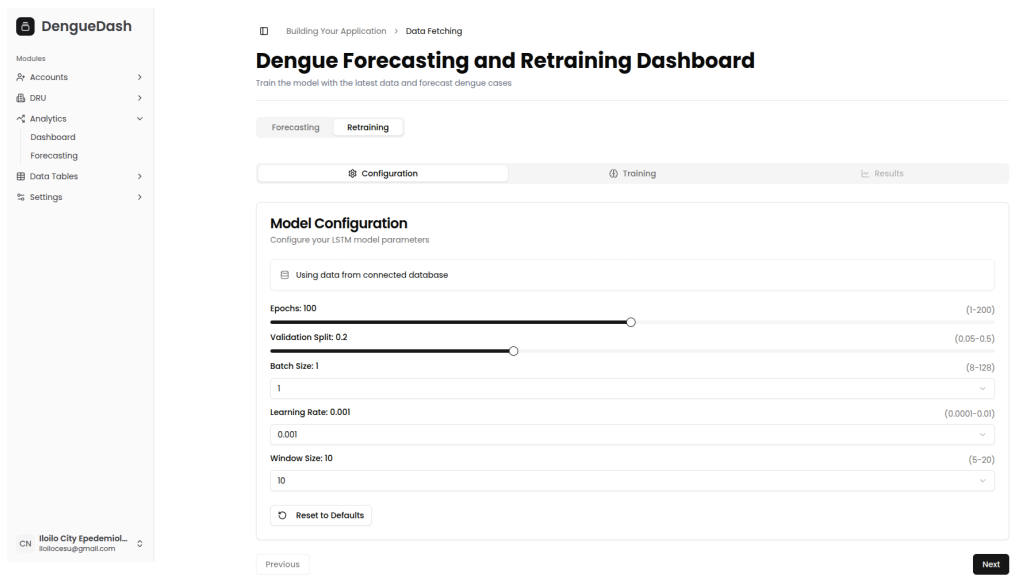


Figure 4.26: Retraining Configurations

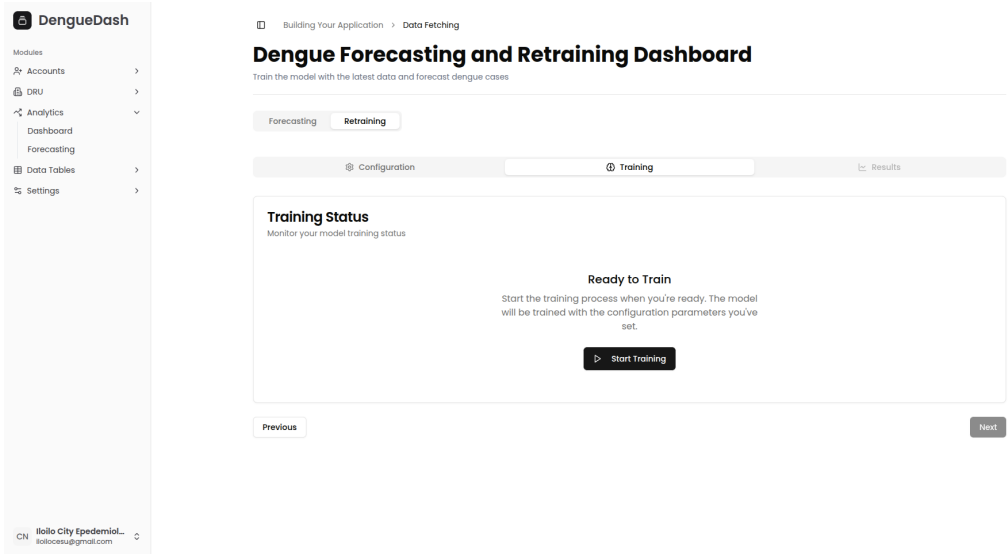


Figure 4.27: Start Retraining

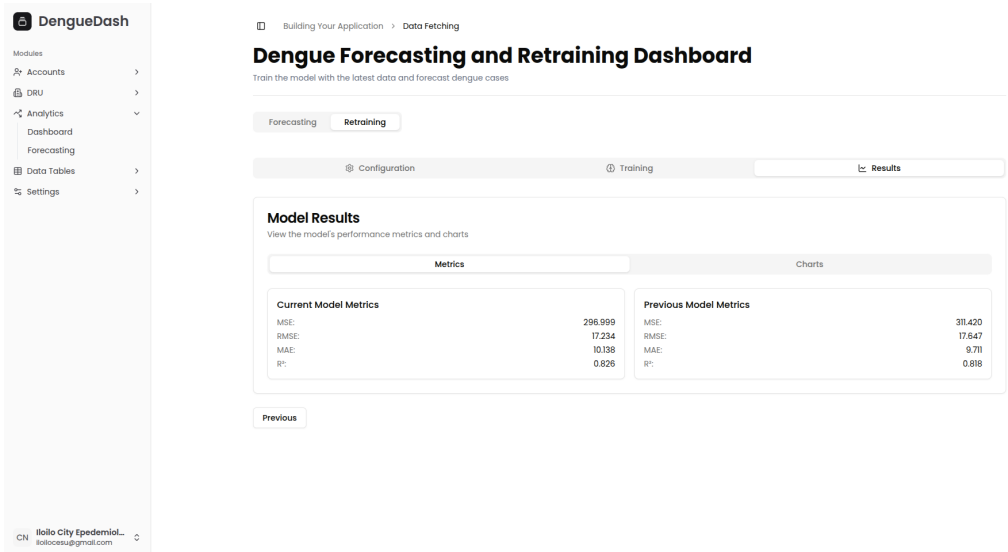


Figure 4.28: Retraining Results

1024 **Managing Accounts**

1025 Proper management of accounts is important to protect the integrity and confi-
1026 dentiality of data. Thus, it is crucial for administrators to track their users and
1027 control the flow of information. As discussed in the user registration of encoders,
1028 admin users from a specific DRU or surveillance have the power to grant them
1029 access to the web application. Figure 4.30 illustrates the interface for this sce-
1030 nario, as the admins can approve or reject their applications. Once approved,
1031 these users can access the features given to encoders and may be promoted to
1032 have administrative access, as shown in Figure 4.31. Otherwise, once deleted, the
1033 email will be blacklisted and is illegible to use when creating another account.
1034 Once deleted, the email will be blacklisted and cannot be used when creating
1035 another account. The same figure also shows the expanded details of the user,
1036 which include personal information and brief activity details within the system.

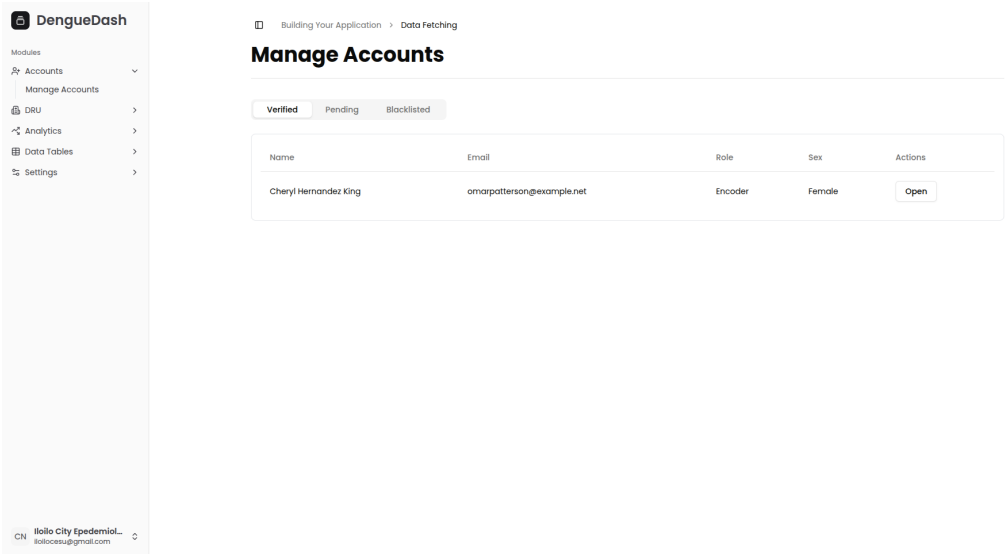


Figure 4.29: List of Verified Accounts

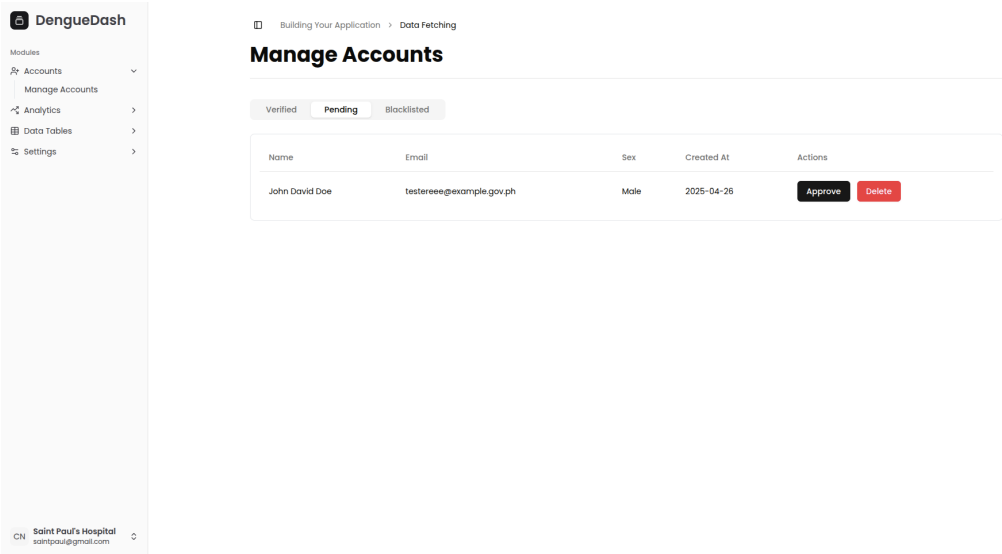


Figure 4.30: List of Pending Accounts

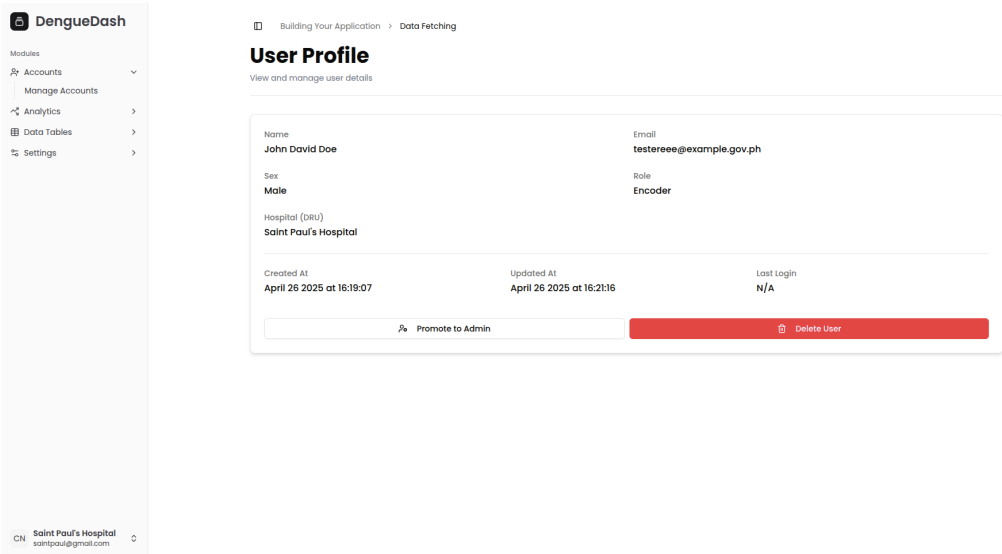


Figure 4.31: Account Details

1037 **Managing DRUs**

1038 Unlike the registration of encoder accounts, the creation of Disease Reporting
1039 Units can only be done within the web application, and the user performing the
1040 creation must be an administrator of a surveillance unit. Figure 4.32 presents the

1041 fields the admin user must fill out, and once completed, the new entry will show
1042 as being managed by that unit, as shown in Figure 4.33. Figure 4.34, on the other
1043 hand, shows the details provided in the registration form as well as its creation
1044 details. There is also an option to delete the DRU, and when invoked, all the
1045 accounts being managed by it, and the cases reported under those accounts will
1046 be deleted.

DengueDash

Modules

- Accounts
- DRU
 - Manage DRU
 - Add DRU
- Analytics
- Data Tables
- Settings

Iloilo City Epidemiol...
ilolocesug@gmail.com

Building Your Application > Data Fetching

Register Disease Reporting Unit

Add a new Disease Reporting Unit to the surveillance system.

Name

Enter DRU name

The official name of the Disease Reporting Unit.

Address Information

Region **Province**

Select Region Select Province

City/Municipality **Barangay**

Select City/Municipality Select Barangay

Street Address

House/Building No, Street Name

Email **Contact Number**

example@health.gov +63 912 345 6789

DRU Type

Select DRU type

The category that best describes this reporting unit.

Register DRU

Figure 4.32: DRU Registration

DengueDash

Modules

- Accounts
- DRU
 - Manage DRU
 - Add DRU
- Analytics
- Data Tables
- Settings

Iloilo City Epidemiol...
ilolocesug@gmail.com

Building Your Application > Data Fetching

Manage Disease Reporting Units

View and manage Disease Reporting Units

DRU Name	Email	
Molo District Health Center	molohealth@gmail.com	Open
Jaro I Health Center	jarohealth@gmail.com	Open
Saint Paul's Hospital	saintpaul@gmail.com	Open

Figure 4.33: List of DRUs

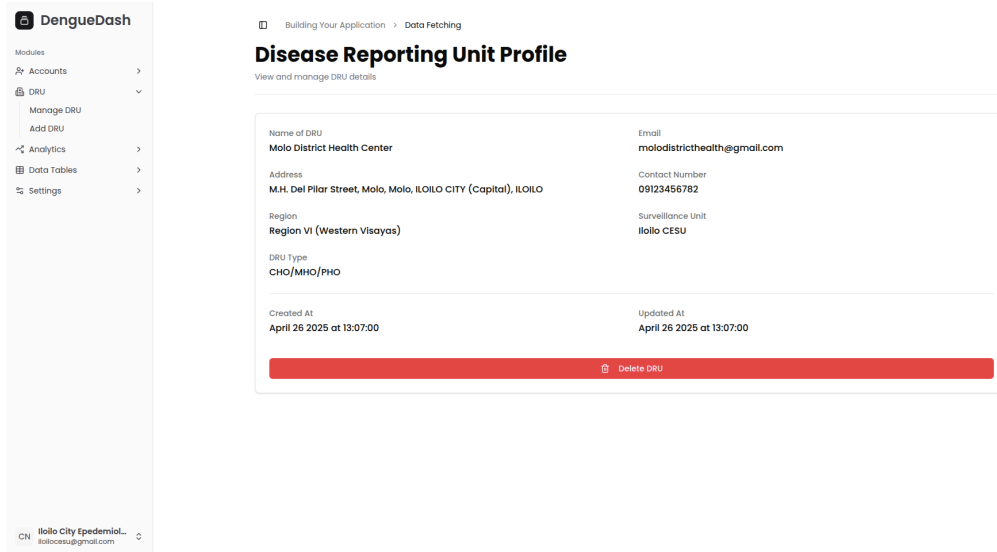


Figure 4.34: DRU details

1047 4.7 User Testing

1048 To evaluate the usability of the system, the System Usability Scale (SUS) was
 1049 utilized. SUS is a Likert-scale-based questionnaire comprising 10 items that are
 1050 critical to assessing system usability. A total of five participants completed the sur-
 1051 vey. Their responses were processed following the step-by-step calculation method
 1052 adopted from (Babich, 2015). The resulting usability scores for each participant
 1053 are shown in Table 4.5.

Participant	Usability Score
1	95.0
2	90.0
3	85.0
4	87.5
5	85.0
Average	88.5

Table 4.5: Computed System Usability Scores per Participant

1054 The average System Usability Scale (SUS) score across systems is typically
 1055 68 (Babich, 2015). In this testing, the system achieved an average SUS score
 1056 of 88.5, indicating a highly positive user experience. This score suggests that
 1057 participants found the system not only enjoyable to use but also intuitive enough

1058 to recommend to others. Furthermore, it demonstrates that the system is suitable
1059 for real-world applications without presenting significant complexity for first-time
1060 users.

References

- 1062 *About GitHub and Git - GitHub Docs.* (n.d.). Retrieved from
1063 [https://docs.github.com/en/get-started/start-your-journey/](https://docs.github.com/en/get-started/start-your-journey/about-github-and-git)
1064 [about-github-and-git](https://docs.github.com/en/get-started/start-your-journey/about-github-and-git)
- 1065 Ahmadini, A. A. H., Naeem, M., Aamir, M., Dewan, R., Alshqaq, S. S. A., &
1066 Mashwani, W. K. (2021). Analysis and forecast of the number of deaths,
1067 recovered cases, and confirmed cases from covid-19 for the top four affected
1068 countries using kalman filter. *Frontiers in Physics*, 9, 629320.
- 1069 Arroyo-Marioli, F., Bullano, F., Kucinskas, S., & Rondón-Moreno, C. (2021).
1070 Tracking r of covid-19: A new real-time estimation using the kalman filter.
1071 *PloS one*, 16(1), e0244474.
- 1072 Babich, N. (2015). *How to use the system usability scale (sus) to*
1073 *evaluate the usability of your website.* Usability Geek. Retrieved
1074 from [https://usabilitygeek.com/how-to-use-the-system-usability](https://usabilitygeek.com/how-to-use-the-system-usability-scale-sus-to-evaluate-the-usability-of-your-website/)
1075 [-scale-sus-to-evaluate-the-usability-of-your-website/](https://usabilitygeek.com/how-to-use-the-system-usability-scale-sus-to-evaluate-the-usability-of-your-website/) (Ac-
1076 cessed: 2025-04-26)
- 1077 Bosano, R. (2023). *Who: Ph most affected by dengue in western pacific.* Retrieved
1078 Use the date of access, from [https://news.abs-cbn.com/spotlight/12/](https://news.abs-cbn.com/spotlight/12/22/23/who-ph-most-affected-by-dengue-in-western-pacific)
1079 [22/23/who-ph-most-affected-by-dengue-in-western-pacific](https://news.abs-cbn.com/spotlight/12/22/23/who-ph-most-affected-by-dengue-in-western-pacific)
- 1080 Brady, O. J., Smith, D. L., Scott, T. W., & Hay, S. I. (2015). Dengue disease
1081 outbreak definitions are implicitly variable. *Epidemics*, 11, 92–102.
- 1082 Bravo, L., Roque, V. G., Brett, J., Dizon, R., & L’Azou, M. (2014). Epidemiology
1083 of dengue disease in the philippines (2000–2011): a systematic literature
1084 review. *PLoS neglected tropical diseases*, 8(11), e3027.
- 1085 Carvajal, T. M., Viacrusis, K. M., Hernandez, L. F. T., Ho, H. T., Amalin, D. M.,
1086 & Watanabe, K. (2018). Machine learning methods reveal the temporal
1087 pattern of dengue incidence using meteorological factors in metropolitan
1088 manila, philippines. *BMC infectious diseases*, 18, 1–15.
- 1089 *Chart.js.* (n.d.). Retrieved from <https://www.chartjs.org/>
- 1090 Cheong, Y. L., Burkart, K., Leitão, P. J., & Lakes, T. (2013). Assessing weather
1091 effects on dengue disease in malaysia. *International journal of environmental*
1092 *research and public health*, 10(12), 6319–6334.

1093 Christie, T. (n.d.). *Home - Django REST framework*. Retrieved from [https://](https://www.django-rest-framework.org/)
1094 www.django-rest-framework.org/

1095 Colón-González, F. J., Fezzi, C., Lake, I. R., & Hunter, P. R. (2013). The effects
1096 of weather and climate change on dengue. *PLoS neglected tropical diseases*,
1097 7(11), e2503.

1098 Hemisphere, N. (2015). Update on the dengue situation in the western pacific
1099 region. *Update*.

1100 Hii, Y. L., Zhu, H., Ng, N., Ng, L. C., & Rocklöv, J. (2012). Forecast of dengue
1101 incidence using temperature and rainfall. *PLoS neglected tropical diseases*,
1102 6(11), e1908.

1103 Joel, C. (2021, 10). *6 reasons to use Tailwind over traditional CSS*. Re-
1104 trieved from [https://dev.to/charliejoel/6-reasons-to-use-tailwind](https://dev.to/charliejoel/6-reasons-to-use-tailwind-over-traditional-css-1nc3)
1105 [-over-traditional-css-1nc3](https://dev.to/charliejoel/6-reasons-to-use-tailwind-over-traditional-css-1nc3)

1106 Leaflet — an open-source JavaScript library for interactive maps. (n.d.). Retrieved
1107 from <https://leafletjs.com/>

1108 Lena, P. (2024). *Iloilo beefs up efforts amid hike in dengue cases*. Retrieved Use
1109 the date of access, from <https://www.pna.gov.ph/articles/1231208>

1110 Li, X., Feng, S., Hou, N., Li, H., Zhang, S., Jian, Z., & Zi, Q. (2022). Applications
1111 of kalman filtering in time series prediction. In *International conference on*
1112 *intelligent robotics and applications* (pp. 520–531).

1113 Ligue, K. D. B., & Ligue, K. J. B. (2022). Deep learning approach to forecasting
1114 dengue cases in davao city using long short-term memory (lstm). *Philippine*
1115 *Journal of Science*, 151(3).

1116 Perla. (2024). *Iloilo beefs up efforts amid hike in dengue cases*. Retrieved Use the
1117 date of access, from <https://www.pna.gov.ph/articles/1231208>

1118 RabDashDC. (2024). *Rabdash dc*. Retrieved Use the date of access, from [https://](https://rabdash.com)
1119 rabdash.com

1120 Runge-Ranzinger, S., Kroeger, A., Oliaro, P., McCall, P. J., Sánchez Tejeda, G.,
1121 Lloyd, L. S., ... Coelho, G. (2016). Dengue contingency planning: from
1122 research to policy and practice. *PLoS neglected tropical diseases*, 10(9),
1123 e0004916.

1124 Shadcn. (n.d.). *Introduction*. Retrieved from <https://ui.shadcn.com/docs>
1125 *Tailwind CSS - Rapidly build modern websites without ever leaving your HTML*.
1126 (n.d.). Retrieved from <https://tailwindcss.com/>

1127 Watts, D. M., Burke, D. S., Harrison, B. A., Whitmire, R. E., & Nisalak, A.
1128 (2020). Effect of temperature on the transmission of dengue virus by aedes
1129 aegypti. *The American Journal of Tropical Medicine and Hygiene*, 36(1),
1130 143–152.

1131 *What is Postman? Postman API Platform*. (n.d.). Retrieved from [https://](https://www.postman.com/product/what-is-postman/)
1132 www.postman.com/product/what-is-postman/

1133 WHO. (2023). *Dengue - global situation*. Retrieved Use the date of ac-
1134 cess, from <https://www.who.int/emergencies/disease-outbreak-news/>

1135 item/2023-DON498
1136 WHO. (2024). *Dengue and severe dengue*. Retrieved Use the date
1137 of access, from [https://www.who.int/news-room/fact-sheets/detail/](https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue)
1138 [dengue-and-severe-dengue](https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue)
1139 *Why Visual Studio Code?* (2021, 11). Retrieved from [https://code](https://code.visualstudio.com/docs/editor/whyvscode)
1140 [.visualstudio.com/docs/editor/whyvscode](https://code.visualstudio.com/docs/editor/whyvscode)
1141 World Health Organization (WHO). (2018). Dengue and severe dengue in the
1142 philippines. *WHO Dengue Factsheet*. (Available at: [https://www.who](https://www.who.int)
1143 [.int](https://www.who.int))
1144 Zhou, S., & Malani, P. (2024). What is dengue? *Jama*, 332(10), 850–850.
1145 Zod. (n.d.). *TypeScript-first schema validation with static type inference*. Re-
1146 trieved from <https://zod.dev/?id=introduction>

1147 **Appendix A**

1148 **Appendix Title**

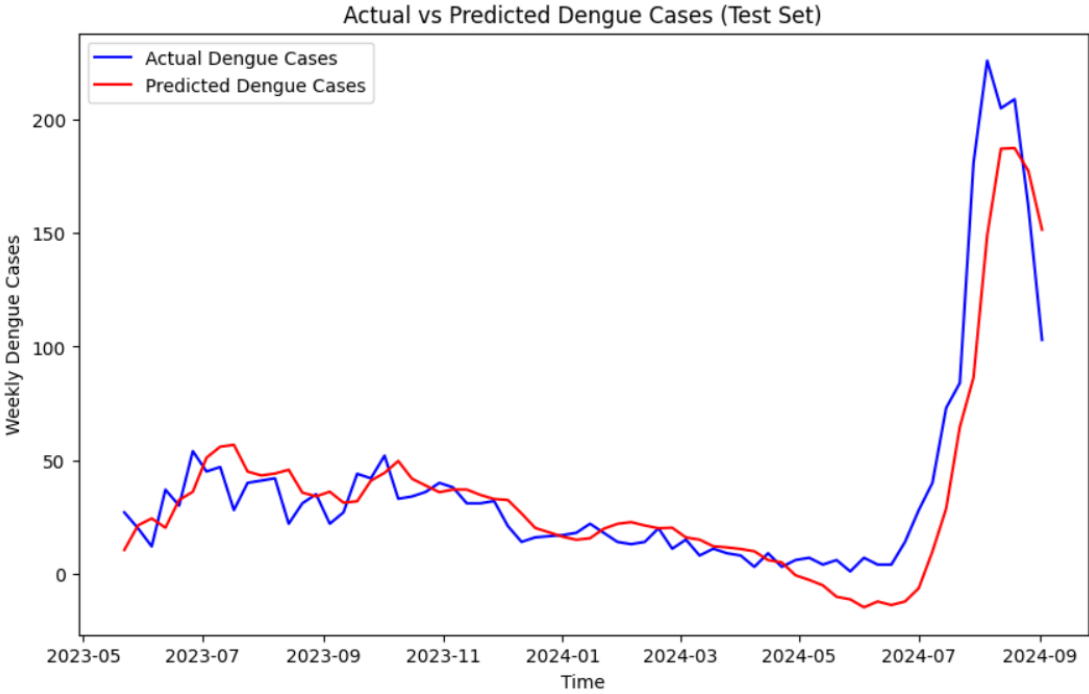


Figure A.1: LSTM Prediction Results for Test Set

1149 **Appendix B**

1150 **Resource Persons**

1151 **Mr. Firstname1 Lastname1**

1152 Role1

1153 Affiliation1

1154 emailaddr1@domain.com

1155 **Ms. Firstname2 Lastname2**

1156 Role2

1157 Affiliation2

1158 emailaddr2@domain.net

1159