Jurnal Teknologi dan Ilmu Komputer Prima (JUTIKOMP) E-ISSN : 2621-234X
Vol. 9 No. 1 April 2026

Development of a Web-Based Diabetes Risk Monitoring System Prototype Using K-Means for Risk Segmentation

Hansen Kusumah Atmaja[a*], **Thommy Willay**[b]
[a,b]*Fakultas Teknologi Informasi, Universitas Widya Dharma Pontianak*
Corresponding Author:
[a]*hansenkusumahatmaja@gmail.com*

ABSTRACT

Diabetes risk monitoring often focuses on data recording and visualization, with limited analytical support for identifying interpretable risk patterns from health-related data. This study developed a web-based diabetes risk monitoring prototype that integrates K-Means clustering for analytical risk segmentation. Using the Pima Indians Diabetes Database with 768 records, this study applied invalid zero handling, median imputation, Min-Max scaling, K-Means clustering, internal validation, cluster profiling, and artifact preparation. Two preprocessing scenarios and candidate values of k were evaluated. The final model was selected under the Extended Scenario at $k = 2$ and produced two analytical risk segments: Higher Risk and Lower Risk. Cluster 0 contained 270 records (35.16%) and was labeled Higher Risk, while Cluster 1 contained 498 records (64.84%) and was labeled Lower Risk. The prototype was implemented using Flask, SQLite, HTML/CSS, Chart.js, and a saved model bundle. It supported data input, result display, prediction history, edit and delete functions, CSV export, and dashboard visualization. Black-box testing across 27 functional test cases showed that all tested features operated successfully. The main contribution is the integration of K-Means-based segmentation into an operational web prototype for structured monitoring, not clinical diagnosis.

Keywords : Diabetes Risk Monitoring, Web-Based Prototype, K-Means Clustering, Risk Segmentation, Health Information System

INTRODUCTION

Diabetes mellitus is one of the most prevalent chronic diseases worldwide and remains a major public health concern because of its increasing prevalence and long-term complications. The International Diabetes Federation estimated that 589 million adults aged 20 to 79 years were living with diabetes in 2024, and this number is projected to reach 853 million by 2050 (International Diabetes Federation, 2025). A large global analysis also reported that diabetes prevalence increased from about 7% in 1990 to nearly 14% in 2022 (Zhou et al., 2024). These trends indicate a growing need for effective monitoring and management strategies. The World Health Organization further noted that many people still lack adequate diabetes monitoring and treatment, which reflects persistent gaps in health services (World Health Organization, 2024).

The same challenge is visible in Indonesia. A 2024 fact sheet issued by Badan Kebijakan Pembangunan Kesehatan, Indonesia's Health Development Policy Agency, reported that the 2023 Indonesian Health Survey showed that 11.7% of the population had indications of diabetes based on blood glucose examination (Badan Kebijakan Pembangunan Kesehatan, 2024). Because diabetes is a chronic disease that can lead to cardiovascular disease, kidney failure, neuropathy, and retinopathy, continuous monitoring of health indicators is essential for early risk identification and prevention.

Although healthcare data have become more widely available, many health information systems still function mainly as data recording and storage tools. They often provide limited analytical support for identifying disease risk patterns. As a result, health data are not fully used to support proactive monitoring. This limitation reduces the ability of healthcare providers and users to interpret health indicators and detect potential risk patterns from available data.

Recent developments in digital health technologies have opened new opportunities for chronic disease monitoring through web-based systems and remote patient monitoring platforms. These systems can improve access to health information, support continuous data collection, and enhance patient engagement (Kerr et al., 2024; Tan et al., 2024). In addition, dashboard-based visualization helps present complex health information in a more interpretable form and supports health-related decision making (Samadbeik et al., 2024). However, many existing systems still emphasize data entry and visualization rather than analytical interpretation.

At the same time, data mining and machine learning methods have been increasingly used in healthcare data analysis. Clustering is one of the most common unsupervised approaches because it can group observations based on similarity without requiring predefined labels. In diabetes-related studies, clustering has been used to identify subgroups with similar metabolic characteristics and risk profiles (Onthoni et al., 2025; Taurbekova et al., 2025). Among clustering methods, K-Means is widely used because it is simple, efficient, and suitable for numerical data.

Despite these developments, previous studies have often treated data analysis and system development separately. Clustering studies usually focus on pattern discovery in datasets without implementing the results in an operational monitoring system. In contrast, web-based monitoring studies often focus on data collection and visualization without integrating analytical models for risk segmentation. This separation reveals a clear research gap.

Therefore, this study aims to develop a web-based diabetes risk monitoring system prototype that integrates the K-Means algorithm to perform analytical risk segmentation based on health-related attributes. The system is intended to support monitoring and interpretation of diabetes-related risk patterns rather than clinical diagnosis. This study contributes in three ways. First, it develops a web-based prototype for structured diabetes risk monitoring. Second, it integrates K-Means clustering to generate data-driven risk segmentation. Third, it presents the clustering results through an interactive dashboard to support more informative health monitoring.

LITERATURE REVIEW

Digital health systems and data-driven analytical methods are both important for improving chronic disease monitoring. This section reviews prior studies on web-based health monitoring systems, diabetes-related health attributes, K-Means clustering for risk segmentation, and evaluation of unsupervised clustering models.

Web-Based Health Monitoring Systems

Web-based health monitoring systems allow users and healthcare providers to access health information, record clinical indicators, and monitor patient conditions through online platforms. In chronic disease management, such systems improve accessibility and support continuous monitoring (Tan et al., 2024). Remote patient monitoring is a key component of this approach because it enables the transmission and observation of physiological data such as blood glucose and blood pressure through digital platforms.

Healthcare dashboards are also important in web-based monitoring environments. Dashboards present clinical data in a visual format that supports interpretation and decision making (Samadbeik et al., 2024). However, many current systems still focus mainly on data entry, storage, and visualization. Their analytical capability for identifying risk patterns remains limited. This indicates that web-based monitoring systems still need stronger integration with analytical models so that stored health data can support not only visualization but also risk-oriented interpretation.

Diabetes Risk Factors and Health Data Attributes

Diabetes mellitus is a metabolic disorder characterized by chronic hyperglycemia caused by insulin resistance or impaired insulin production. Several physiological variables are strongly associated with diabetes risk, including blood glucose, body mass index, age, blood pressure, and insulin level (Zhou et al., 2024). Blood glucose is a central indicator because it directly reflects impaired glucose metabolism. BMI is also important because obesity is closely associated with insulin resistance and diabetes risk. Age and blood pressure are additional risk-related variables because metabolic function often declines with age, while hypertension is commonly associated with metabolic and cardiovascular disorders (Taurbekova et al., 2025). Public diabetes datasets also often include skin thickness and diabetes pedigree function. These variables can support a more detailed description of body composition and hereditary susceptibility in analytical studies (Onthoni et al., 2025).

K-Means Clustering for Risk Segmentation

Clustering is an unsupervised learning technique used to group observations with similar characteristics. In healthcare research, clustering is useful for exploring hidden structures in patient data and identifying subgroups with different profiles (Taurbekova et al., 2025). K-Means is one of the most widely used clustering algorithms because it is simple, efficient, and scalable for numerical datasets.

K-Means partitions observations into a predefined number of clusters and assigns each observation to the nearest centroid using a distance-based approach. The centroids are then updated iteratively until convergence is achieved (scikit-learn developers, 2026c). In diabetes research, K-Means has been used to identify subgroups with similar risk patterns and metabolic characteristics (Onthoni et al., 2025). However, the method is sensitive to centroid initialization and requires prior specification of the number of clusters. In addition, many clustering studies remain dataset-oriented and do not extend the identified segments into a working monitoring prototype. This creates an opportunity to connect clustering results with operational web-based monitoring functions.

Evaluation of Unsupervised Clustering Models

Because clustering is performed without labeled targets, cluster quality must be evaluated using internal validation metrics. Commonly used indices include the Silhouette Coefficient, Calinski-Harabasz Index, and Davies-Bouldin Index. The Silhouette Coefficient measures how well an observation fits within its assigned cluster compared with other clusters. Higher values indicate better separation. The Calinski-Harabasz Index evaluates the ratio of between-cluster dispersion to within-cluster dispersion, where higher values indicate better clustering quality. The Davies-Bouldin Index measures the similarity among clusters, where lower values indicate better separation (Chicco et al., 2025; Ikotun et al., 2025).

Using multiple internal validity indices provides a more reliable basis for selecting the best clustering configuration because different metrics may respond differently to the same data structure.

Research Gap

Prior studies show that digital health monitoring systems and clustering-based health data analysis are both useful, but they are often developed separately. Web-based monitoring studies mainly emphasize data collection and visualization, whereas clustering studies focus on analytical exploration of datasets. The integration of clustering-based analytical techniques into a prototype-level web-based monitoring system remains limited. This study addresses that gap by combining K-Means-based risk segmentation with a web-based diabetes monitoring prototype that can present segmentation results through history records, CSV export, and dashboard visualization.

METHODS

Research Design and Sampling

This study adopted an applied quantitative design to develop a web-based prototype for diabetes risk monitoring through data-driven analytical risk segmentation. The analytical component used K-Means clustering to group observations according to similarity in health-related attributes. The study did not aim to produce a clinical diagnosis. It focused on generating analytical risk segments that could support monitoring. The study used the Pima Indians Diabetes Database as a public secondary dataset. The dataset contained 768 records and 9 variables, consisting of 8 health-related predictor attributes and 1 outcome variable. Because the dataset was fully accessible and manageable in size, all available records were included after screening and preprocessing. Accordingly, the clustering experiment used the complete dataset rather than a sampled subset (Chang et al., 2023; Feng et al., 2023).

Data Collection

Data collection was conducted through secondary data acquisition. The dataset was prepared in CSV format and processed in a reproducible Python environment developed in Visual

Studio Code. The analytical workflow covered dataset inspection, preprocessing, clustering, internal validation, cluster profiling, and preparation of reusable artifacts for prototype integration. To structure this workflow, the study followed the general logic of CRISP-DM, particularly the phases of data understanding, data preparation, modeling, evaluation, and deployment-oriented preparation. Figure 1 shows the research workflow and analytical pipeline, which consisted of public dataset acquisition, dataset inspection, invalid zero handling, median imputation, Min-Max scaling, K-Means clustering, internal validation, cluster profiling, and artifact preparation (Plotnikova et al., 2022; Schröer et al., 2021).

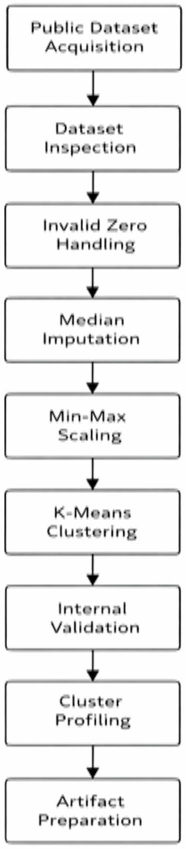


Figure 1. Research Workflow / Analytical Pipeline

Measures

Eight numerical variables were used as clustering features: Pregnancies, Glucose, BloodPressure, SkinThickness, Insulin, BMI, DiabetesPedigreeFunction, and Age. These variables were selected because they represent health-related indicators available in the dataset and were directly used as analytical inputs. The Outcome variable was not used as an input for clustering because K-Means is an unsupervised method. Outcome was retained only for post-clustering profiling and interpretation. Unit labels were used consistently in the prototype interface to support interpretation of numerical health attributes. Pregnancies was presented as a count variable, Glucose in mg/dL, BloodPressure in mm Hg, SkinThickness in mm, Insulin in µU/mL, BMI in kg/m², DiabetesPedigreeFunction as a score, and Age in years. Table 1 summarizes the variables used in the clustering analysis and their roles in the study (Chang et al., 2023; Feng et al., 2023).

Table 1. Variables Used in the Clustering Analysis

Variable	Description	Unit/Scale	Data Type	Role in Analysis
Pregnancies	Number of previous pregnancies	Count	Numeric	Input feature
Glucose	Plasma glucose concentration	mg/dL	Numeric	Input feature
BloodPressure	Diastolic blood pressure	mm Hg	Numeric	Input feature
SkinThickness	Triceps skin fold thickness	mm	Numeric	Input feature
Insulin	Two-hour serum insulin	µU/mL	Numeric	Input feature
BMI	Body mass index	kg/m²	Numeric	Input feature
DiabetesPedigreeFunction	Diabetes pedigree score	score	Numeric	Input feature
Age	Age	years	Numeric	Input feature
Outcome	Diabetes outcome label in the original dataset	0 or 1	Binary numeric	Post-clustering profiling only

Data Preprocessing

The preprocessing stage began with inspection of dataset structure, variable types, descriptive statistics, missing values, and duplicate records. The study then compared two preprocessing scenarios. The Baseline Scenario treated zero values in Glucose, BloodPressure, and BMI as invalid. The Extended Scenario additionally treated zero values in SkinThickness and Insulin as invalid. This comparison was performed to determine the most appropriate preparation strategy before final model selection.

In the final preprocessing workflow, zero values in Glucose, BloodPressure, SkinThickness, Insulin, and BMI were recoded as missing values because they were considered clinically implausible in this analytical context. Missing values were then handled using median imputation. After imputation, all clustering features were normalized using Min-Max scaling to a comparable range between 0 and 1. This step was necessary because K-Means is distance-based and sensitive to differences in variable scale. Table 2 presents the preprocessing rules applied in the final workflow (Abousaber et al., 2025; scikit-learn developers, 2026g; scikit-learn developers, 2026d; Feng et al., 2023).

Table 2. Preprocessing Rules Applied in the Final Workflow

Step	Variables Affected	Technique Applied	Purpose
Initial inspection	All variables	Structure, type, descriptive, missing value, and duplicate checking	To understand dataset quality before modeling
Invalid zero handling	Glucose, BloodPressure, SkinThickness, Insulin, BMI	Zero values recoded as NaN	To treat clinically implausible zero values as hidden missing values
Missing value handling	Glucose, BloodPressure, SkinThickness, Insulin, BMI	Median imputation	To replace missing values with robust central tendency estimates
Feature scaling	All clustering features	Min-Max Scaling	To normalize the variable range before distance-based clustering
Feature selection	Eight health attributes	Inclusion as clustering input	To define the final analytical features for K-Means
Outcome handling	Outcome	Excluded from clustering, retained for profiling	To support cluster interpretation without affecting unsupervised clustering

K-Means Clustering Procedure

The clustering stage was implemented using the K-Means algorithm from scikit-learn. K-Means was selected because it is computationally efficient and suitable for grouping observations with similar numerical characteristics. The model was configured with k-means++ initialization, n_init = "auto", max_iter = 300, and random_state = 42 to support stable and reproducible clustering. To determine the most appropriate number of clusters, the study evaluated candidate values of $k = 2, 3, 4, 5$, and 6 after preprocessing (scikit-learn developers, 2026c).

Cluster Evaluation and Model Selection

Because the clustering task did not rely on predefined class labels, model quality was assessed using three internal validation metrics: Silhouette Coefficient, Calinski-Harabasz Index, and Davies-Bouldin Index. Higher values indicate better clustering quality for Silhouette and Calinski-Harabasz, whereas lower values indicate better clustering quality for Davies-Bouldin. To support joint comparison across these metrics, each preprocessing scenario and candidate *k* combination was ranked on each metric, and the ranks were summed as Total Rank. A lower Total Rank indicated better overall clustering performance. The final analytical model was selected by jointly comparing clustering results across preprocessing scenarios and candidate *k* values. This approach reduced dependence on a single validity index and provided a broader basis for selecting the final configuration (Chicco et al., 2025; scikit-learn developers, 2026f; scikit-learn developers, 2026a; scikit-learn developers, 2026b; Ikotun et al., 2025).

Cluster Profiling and Risk Interpretation

After the final clustering configuration was selected, each cluster was profiled using cluster size, mean values of the original variables, and the proportion of Outcome = 1 in each cluster. This profiling stage was used to interpret the substantive meaning of the resulting groups. A cluster with higher mean values in key variables, especially Glucose, BMI, and Age, together with a higher proportion of Outcome = 1, was interpreted as a Higher Risk segment. A cluster with lower mean values and a lower proportion of Outcome = 1 was interpreted as a Lower Risk segment. These labels were used strictly for analytical risk segmentation and not for clinical diagnosis (Chang et al., 2023; Feng et al., 2023).

Prototype-Oriented Artifact Preparation and System Design

To support integration into the web-based prototype, the final preprocessing object, K-Means model, cluster-label mapping, and model metadata were saved as reusable artifacts. This artifact preparation ensured that the same preprocessing logic and clustering configuration could be applied consistently to new user input (scikit-learn developers, 2026e).

The prototype was designed using four main components: web interface, backend application, clustering engine, and database. In the implemented workflow, user-entered health data were validated by the backend, transformed using the saved preprocessing object, assigned to a cluster using the final K-Means model, and returned as an analytical risk segmentation result for monitoring and dashboard visualization. The backend was implemented using Flask, prediction records were stored in a local SQLite database, model integration used final_model_bundle.joblib, and the interface was built using HTML templates, CSS, and Chart.js. Unit labels were displayed consistently in the input, edit, result, history, and dashboard pages to support interpretation of numerical attributes. Pregnancies was displayed as a count variable, while Glucose, BloodPressure, SkinThickness, Insulin, BMI, DiabetesPedigreeFunction, and Age were displayed with their respective units.

UML-based modeling was used to document the prototype design. Figure 2 presents the use case diagram, which includes one actor, User, and eight main functions: Input Health Data, Process Analytical Risk Segmentation, View Result, View History, Edit Record, Delete Record, Download CSV, and View Dashboard. Figure 3 presents the activity diagram, which shows the workflow from data entry and input validation to preprocessing, K-Means prediction, risk-label mapping, database storage, and result display. The invalid-input branch returns the user to the input form, while the valid-input branch completes the analytical risk segmentation workflow.

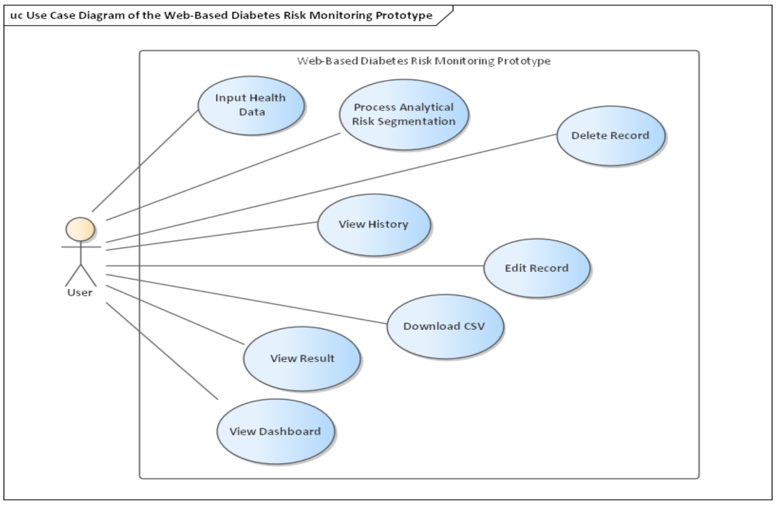


Figure 2. Use Case Diagram of the Web-Based Diabetes Risk Monitoring Prototype

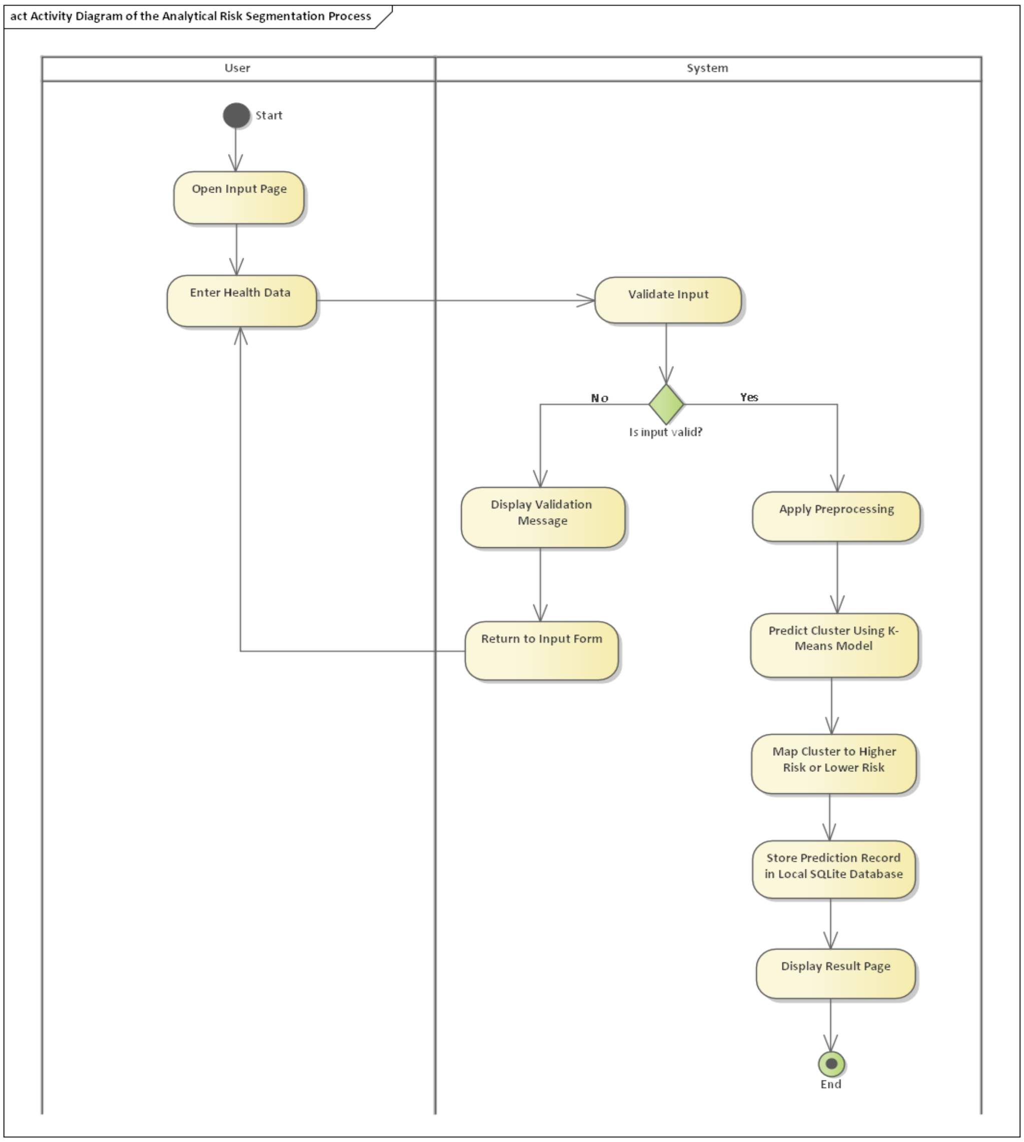


Figure 3. Activity Diagram of the Analytical Risk Segmentation Process

RESULTS

Dataset Description

The Pima Indians Diabetes Database was used as the primary dataset for clustering analysis. It contained 768 records and 9 variables, consisting of 8 health-related attributes used as clustering features and 1 additional variable, Outcome, which was retained only for post-clustering profiling. The clustering features were Pregnancies, Glucose, BloodPressure, SkinThickness, Insulin, BMI, DiabetesPedigreeFunction, and Age. Table 3 summarizes the dataset structure.

Table 3. Dataset Summary

Item	Value
Number of records	768
Number of variables	9
Clustering features	8
Additional profiling variable	Outcome

Initial inspection showed heterogeneous variable scales and several zero values in clinical attributes, indicating the need for preprocessing before clustering. The descriptive statistics showed mean values of 120.895 for Glucose (mg/dL), 69.105 for BloodPressure (mm Hg), 31.993 for BMI (kg/m²), and 33.241 for Age (years). The mean value of Outcome was 0.349, indicating that 34.9% of the records were associated with Outcome = 1 in the original dataset.

Preprocessing Results

The preprocessing stage included inspection of data structure, descriptive statistics, missing values, and duplicate records. No duplicate records were found. Two preprocessing scenarios were compared. The Baseline Scenario treated zero values in Glucose, BloodPressure, and BMI as invalid, whereas the Extended Scenario additionally treated zero values in SkinThickness and Insulin as invalid. In the final workflow, zero values in Glucose, BloodPressure, SkinThickness, Insulin, and BMI were recoded as missing values, with counts of 5, 35, 227, 374, and 11, respectively. These values were imputed using the median of each variable. After imputation, no missing values remained in the clustering features, and all features were normalized using Min-Max Scaling before K-Means clustering.

Results of Cluster Evaluation and Model Selection

K-Means was evaluated under two preprocessing scenarios: the Baseline Scenario and the Extended Scenario. The Baseline Scenario treated zero values in Glucose, BloodPressure, and BMI as invalid, whereas the Extended Scenario additionally treated zero values in SkinThickness and Insulin as invalid. Candidate values of $k = 2, 3, 4, 5$, and 6 were evaluated using three internal validation metrics: Silhouette Score, Calinski-Harabasz Index, and Davies-Bouldin Index. Table 4 presents the complete evaluation results.

Table 4. Cluster Evaluation Results Across Preprocessing Scenarios and Candidate k Values

Scenario	*k*	Silhouette Score	Calinski-Harabasz Index	Davies-Bouldin Index	Total Rank
Baseline Scenario	2	0.2475	230.7728	1.6945	4.0
Baseline Scenario	3	0.2173	197.1482	1.6291	5.0
Baseline Scenario	4	0.1511	161.4695	1.7865	9.0
Baseline Scenario	5	0.1266	138.5570	1.9910	14.0
Baseline Scenario	6	0.1398	130.0772	1.9016	13.0
Extended Scenario	**2**	**0.2731**	**271.5499**	**1.5386**	**3.0**
Extended Scenario	3	0.2378	224.3204	1.5936	6.0
Extended Scenario	4	0.1968	186.3429	1.7430	10.0
Extended Scenario	5	0.1673	154.9913	1.7557	13.0
Extended Scenario	6	0.1581	148.7285	1.7156	13.0

Notes:
Higher Silhouette Score and Calinski-Harabasz Index values indicate better clustering performance, whereas lower Davies-Bouldin Index values indicate better clustering performance. Total Rank was calculated separately within each preprocessing scenario by ranking candidate *k* values across the three internal validation metrics and then summing the ranks. Silhouette Score and Calinski-Harabasz Index were ranked in descending order, while Davies-Bouldin Index was ranked in ascending order. A lower Total Rank indicates better overall clustering performance within the same preprocessing scenario.

The final model was selected by comparing the best-ranked configuration from each preprocessing scenario and by considering the absolute values of the three validation metrics.

Under the Baseline Scenario, the best result was obtained at $k = 2$, with a Silhouette Score of 0.2475, a Calinski-Harabasz Index of 230.7728, a Davies-Bouldin Index of 1.6945, and a Total Rank of 4.0. Under the Extended Scenario, the best result was also obtained at $k = 2$, with a Silhouette Score of 0.2731, a Calinski-Harabasz Index of 271.5499, a Davies-Bouldin Index of 1.5386, and a Total Rank of 3.0. Compared with the best Baseline Scenario configuration, the Extended Scenario at $k = 2$ produced a higher Silhouette Score, a higher Calinski-Harabasz Index, and a lower Davies-Bouldin Index. Therefore, this configuration was selected as the final analytical model for diabetes risk segmentation. Figure 4 shows the silhouette comparison across candidate *k* values. The silhouette scores generally decreased after k = 2, indicating that the two-cluster solution provided the strongest cluster separation across both preprocessing scenarios.

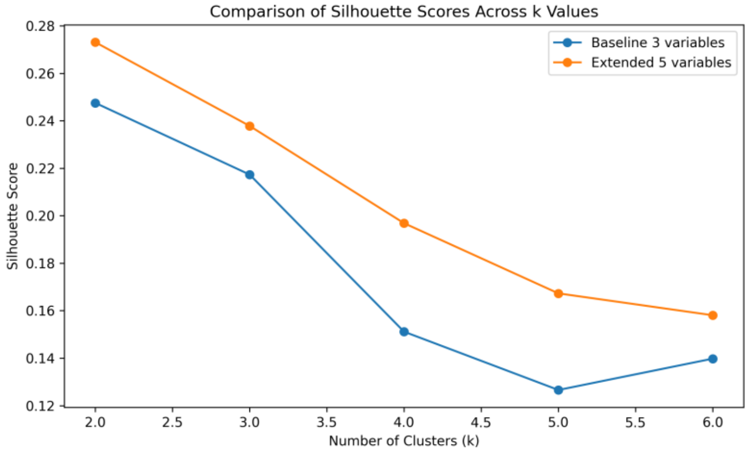


Figure 4. Comparison of Silhouette Scores Across k Values

Final Risk Segmentation Results

Using the selected Extended Scenario at *k* = 2, the final model produced two clusters that were used as analytical risk segments. Cluster 0 contained 270 records, or 35.16% of the dataset, while Cluster 1 contained 498 records, or 64.84%.

Cluster 0 showed higher mean values for Glucose (134.915 mg/dL), BMI (33.189 kg/m^2), and Age (45.748 years), with Outcome = 1 reaching 54.44%. Cluster 1 showed lower mean values for Glucose (114.468 mg/dL), BMI (32.058 kg/m^2), and Age (26.460 years), with Outcome = 1 at 24.30%. Based on this profile, Cluster 0 was labeled Higher Risk and Cluster 1 was labeled Lower Risk. Table 5 summarizes the final cluster profile, and Figure 5 shows the final cluster distribution.

Table 5. Final Cluster Profile and Risk Interpretation

Cluster	Number of Records	Percentage (%)	Mean Glucose (mg/dL)	Mean BMI (kg/m^2)	Mean Age (years)	Outcome = 1 (%)	Interpretation
0	270	35.16	134.915	33.189	45.748	54.44	Higher Risk
1	498	64.84	114.468	32.058	26.460	24.30	Lower Risk

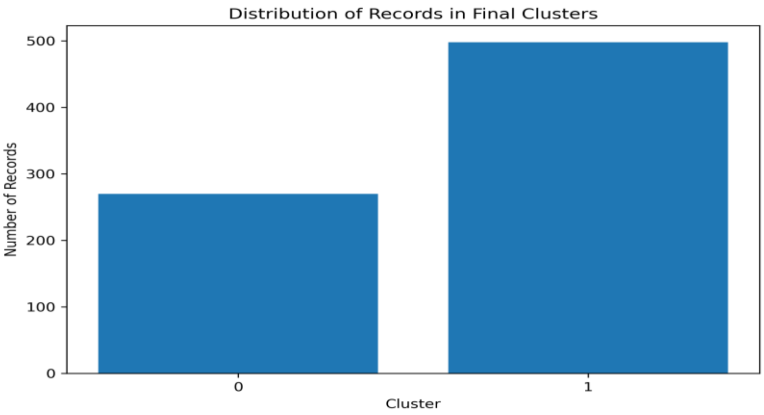


Figure 5. Distribution of Records in Final Clusters

Implementation of the Web-Based Prototype

The web-based prototype was implemented as a Flask application with a local SQLite database. The backend loaded final_model_bundle.joblib, which contained the features list, invalid_zero_cols, a median imputer, a Min-Max scaler, a trained K-Means model with *k* = 2, and the cluster label mapping for Higher Risk and Lower Risk. The interface used HTML templates and CSS, while dashboard charts were rendered with Chart.js. Figure 6 shows the final system architecture, and Table 6 summarizes the implemented system modules and routes.

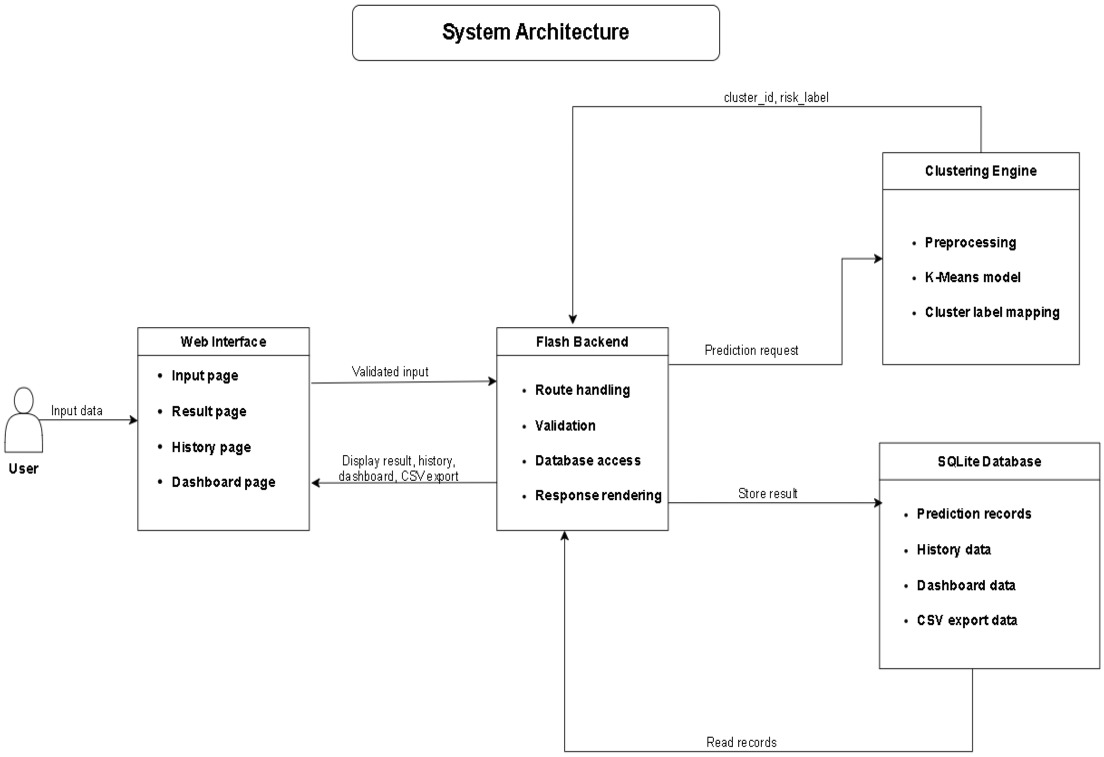


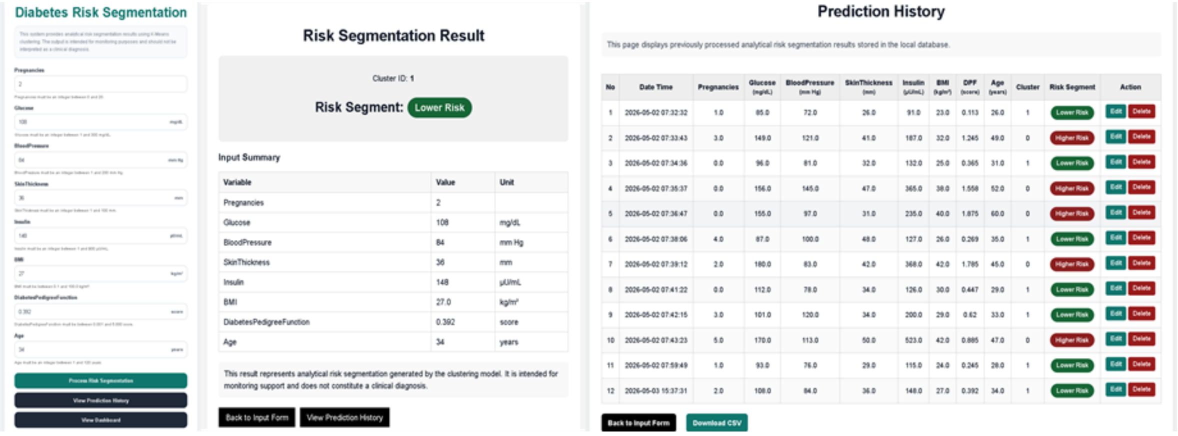
Figure 6. Final System Architecture of the Web-Based Analytical Risk Segmentation Prototype

Table 6. System Modules and Routes of the Implemented Web-Based Prototype

Module/Page	Main Function	Input	Output	Route
Input Page	Receives health data	Eight input features	Validated input with unit indicators	GET /
Result Page	Displays segmentation result	Validated input, predicted cluster	Cluster ID, badge label, input summary, units	POST /predict
History Page	Displays stored records	Prediction records	History table, unit labels, badge labels, edit, delete, CSV export	GET /history
Edit Function	Updates stored record	Record ID, edited inputs	Updated cluster and risk label	GET /history/edit/<id>, POST /history/update/<id>
Delete Function	Confirms deletion of selected history record	Selected record from history table	Custom confirmation modal, confirmed deletion or cancellation, updated history	POST /history/delete/<id>
Dashboard Page	Shows operational summary	Aggregated records	Summary cards, charts with unit labels, recent history (max 10), badge labels	GET /dashboard
CSV Export	Exports history data	Prediction records	prediction_history.csv	GET /history/download

The main input page was served through GET /, while valid submissions were processed through POST /predict. This route validated the submitted input, applied preprocessing, predicted the cluster, mapped the cluster to an analytical risk segmentation label, stored the processed record in the local SQLite database, and displayed the result page. The input and edit pages displayed unit indicators for Glucose (mg/dL), BloodPressure (mm Hg), SkinThickness (mm), Insulin (µU/mL), BMI (kg/m²), DiabetesPedigreeFunction (score), and Age (years), while Pregnancies was displayed without a measurement unit. The result, history, and dashboard pages used badge-based visual labels for Higher Risk and Lower Risk to maintain interface consistency. The displayed units served as interface-level labels to improve readability of health-related inputs and stored prediction records. The analytical process still used numerical values after preprocessing and scaling before K-Means assignment.

The history module allowed users to view stored records, edit records, delete records, and export prediction history as prediction_history.csv. Edited records were revalidated and reprocessed using the same K-Means model before being updated in the local database. The dashboard displayed summary cards, a doughnut chart for risk segment distribution, a bar chart comparing mean Glucose (mg/dL), BMI (kg/m²), and Age (years), and a recent prediction history panel limited to the latest 10 records. The dashboard presented current local prediction records processed through the prototype rather than the raw research dataset. Figure 7 shows the main input page, result page, and history page, while Figure 8 shows the dashboard page.



Risk Segmentation Result

Cluster ID: 1

Risk Segment: Lower Risk

Input Summary

Variable	Value	Unit
Pregnancies	2	
Glucose	108	mg/dL
BloodPressure	84	mm Hg
SkinThickness	36	mm
Insulin	148	µU/mL
BMI	27.0	kg/m²
DiabetesPedigreeFunction	0.392	score
Age	34	years

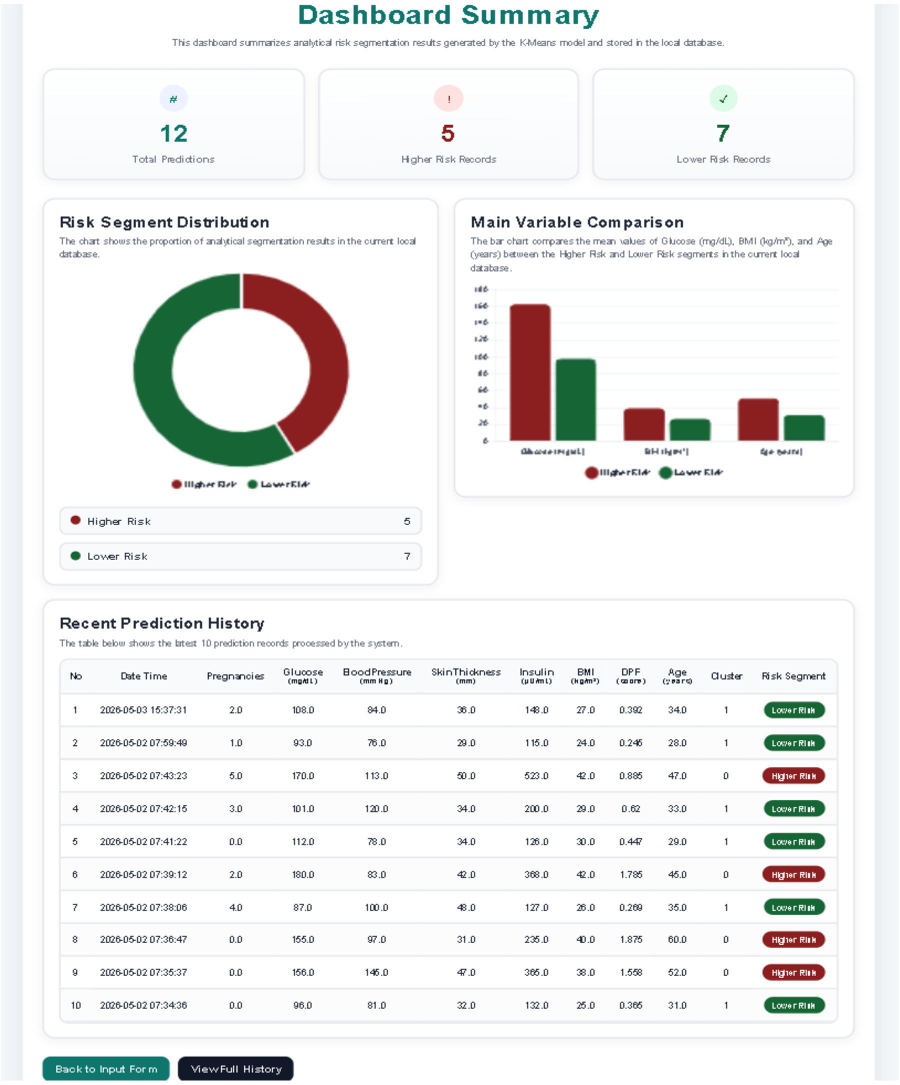
This result represents analytical risk segmentation generated by the clustering model. It is intended for monitoring support and does not constitute a clinical diagnosis.

Prediction History

This page displays previously processed analytical risk segmentation results stored in the local database.

No	Date Time	Pregnancies	Glucose (mg/dL)	BloodPressure (mm Hg)	SkinThickness (mm)	Insulin (µU/mL)	BMI (kg/m²)	DPF (score)	Age (years)	Cluster	Risk Segment	Action
1	2026-05-02 07:32:32	1.0	85.0	72.0	26.0	91.0	23.0	0.113	26.0	1	Lower Risk	Edit Delete
2	2026-05-02 07:33:43	3.0	149.0	121.0	41.0	187.0	32.0	1.245	49.0	0	Higher Risk	Edit Delete
3	2026-05-02 07:34:36	0.0	96.0	81.0	32.0	132.0	25.0	0.365	31.0	1	Lower Risk	Edit Delete
4	2026-05-02 07:35:37	0.0	156.0	145.0	47.0	365.0	38.0	1.558	52.0	0	Higher Risk	Edit Delete
5	2026-05-02 07:36:47	0.0	155.0	97.0	31.0	235.0	40.0	1.875	60.0	0	Higher Risk	Edit Delete
6	2026-05-02 07:38:06	4.0	87.0	100.0	48.0	127.0	26.0	0.269	35.0	1	Lower Risk	Edit Delete
7	2026-05-02 07:39:12	2.0	180.0	83.0	42.0	368.0	42.0	1.785	45.0	0	Higher Risk	Edit Delete
8	2026-05-02 07:41:22	0.0	112.0	78.0	34.0	126.0	30.0	0.447	29.0	1	Lower Risk	Edit Delete
9	2026-05-02 07:42:15	3.0	101.0	120.0	34.0	200.0	29.0	0.62	33.0	1	Lower Risk	Edit Delete
10	2026-05-02 07:43:23	5.0	170.0	113.0	50.0	523.0	42.0	0.885	47.0	0	Higher Risk	Edit Delete
11	2026-05-02 07:59:49	1.0	93.0	76.0	29.0	115.0	24.0	0.245	28.0	1	Lower Risk	Edit Delete
12	2026-05-03 15:37:31	2.0	108.0	84.0	36.0	148.0	27.0	0.392	34.0	1	Lower Risk	Edit Delete

Figure 7. Core Interface Pages of the Web-Based Diabetes Risk Monitoring Prototype: (a) main input page with numeric validation and unit display, (b) result page showing cluster ID, badge-based risk label, input summary, and units, and (c) history page showing stored records with unit labels, badge-based risk labels, edit, delete, and CSV export actions.



Dashboard Summary

This dashboard summarizes analytical risk segmentation results generated by the K-Means model and stored in the local database.

12 Total Predictions

5 Higher Risk Records

7 Lower Risk Records

Risk Segment Distribution

The chart shows the proportion of analytical segmentation results in the current local database.

● Higher Risk	5
● Lower Risk	7

Main Variable Comparison

The bar chart compares the mean values of Glucose (mg/dL), BMI (kg/m²), and Age (years) between the Higher Risk and Lower Risk segments in the current local database.

Recent Prediction History

The table below shows the latest 10 prediction records processed by the system.

No	Date Time	Pregnancies	Glucose (mg/dL)	BloodPressure (mm Hg)	SkinThickness (mm)	Insulin (μU/mL)	BMI (kg/m²)	DPF (score)	Age (years)	Cluster	Risk Segment
1	2026-05-03 15:37:31	2.0	108.0	84.0	36.0	148.0	27.0	0.392	34.0	1	Lower Risk
2	2026-05-02 07:59:49	1.0	93.0	76.0	29.0	115.0	24.0	0.246	28.0	1	Lower Risk
3	2026-05-02 07:43:23	5.0	170.0	113.0	50.0	523.0	42.0	0.885	47.0	0	Higher Risk
4	2026-05-02 07:42:15	3.0	101.0	120.0	34.0	200.0	29.0	0.62	33.0	1	Lower Risk
5	2026-05-02 07:41:22	0.0	112.0	78.0	34.0	126.0	30.0	0.447	29.0	1	Lower Risk
6	2026-05-02 07:39:12	2.0	180.0	83.0	42.0	368.0	42.0	1.785	45.0	0	Higher Risk
7	2026-05-02 07:38:06	4.0	87.0	100.0	48.0	127.0	26.0	0.269	35.0	1	Lower Risk
8	2026-05-02 07:36:47	0.0	155.0	97.0	31.0	235.0	40.0	1.875	60.0	0	Higher Risk
9	2026-05-02 07:35:37	0.0	156.0	146.0	47.0	365.0	38.0	1.558	52.0	0	Higher Risk
10	2026-05-02 07:34:36	0.0	96.0	81.0	32.0	132.0	25.0	0.365	31.0	1	Lower Risk

Figure 8. Dashboard Page Showing Operational Summaries, Risk Distribution, Main Variable Comparison with Unit Labels, and Recent Prediction History with Badge-Based Risk Labels

System Testing Results

System testing used manual black-box functional testing. The prototype was tested as a local Flask web application running at http://127.0.0.1:5000. The testing environment used Python 3.14.3, Flask 3.1.3, Microsoft Edge, Windows 11, a local SQLite database, and curl for selected route-level checks.

A total of 27 functional test cases were executed, and all 27 passed. The validation mechanism was designed as technical range-based validation for prototype operation, not as clinical plausibility validation. The test cases covered page accessibility and navigation, client-side validation, server-side validation for bypass cases, prediction storage, CRUD operations in the history module, CSV export, and dashboard consistency. Table 7 summarizes the grouped testing results. The main input page, result page, history page, and dashboard page were reachable and responded according to their implemented routes. Valid processed input was also stored correctly in the local SQLite database and appeared in the history table.

Table 7. Summary of Black-Box Functional Testing Results for the Web-Based Prototype

Functional Group	Scope of Verified Functions	Result
Page accessibility and navigation	Main page, result page, history page, dashboard page, navigation buttons	Passed
Client-side validation	Required fields, out-of-range values, decimal Age input	Passed
Server-side validation	Bypass cases submitted through POST /predict	Passed
Prediction storage	Saving valid processed input into local SQLite database	Passed
CRUD operations in history	Edit, update, custom delete confirmation modal, cancellation without deletion, confirmed deletion, non-existent record handling	Passed
CSV export	Download of prediction_history.csv	Passed
Dashboard consistency	Summary and recent history after update and delete	Passed

The delete workflow used a custom confirmation modal that displayed selected record details before the user either cancelled the action or confirmed deletion. Decimal Age input, specifically Age = 30.5, was blocked by browser validation and prevented form submission, while invalid bypass cases submitted through POST /predict were rejected and did not create new prediction records. The edit function accepted valid updates and recalculated analytical risk segmentation. The delete workflow also allowed cancellation without removing the selected record, deleted the selected record after confirmation, and returned “Record not found.” when a non-existent record ID was submitted. The CSV export function generated and downloaded prediction_history.csv successfully. The dashboard summary cards and recent prediction history reflected the latest contents of the local SQLite database after update and delete operations. Figure 9a and figure 9b shows representative testing evidence.

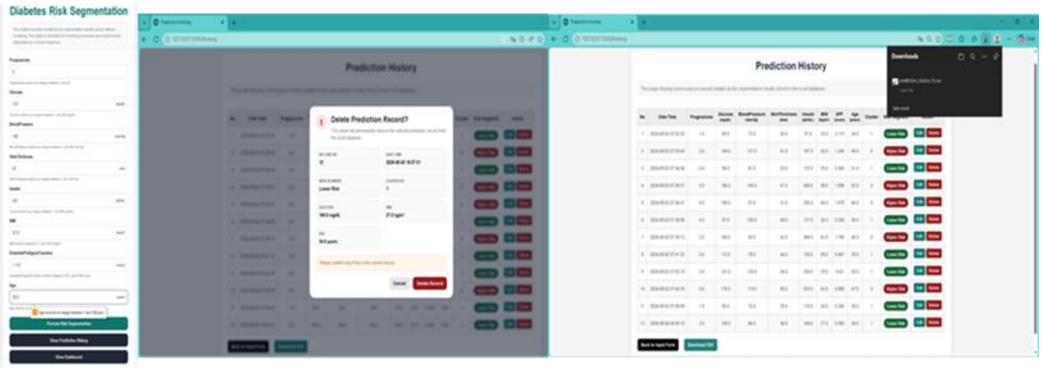


Figure 9a. Representative Black-Box Functional Testing Evidence: (a) range-based validation for decimal Age input, (b) custom delete confirmation modal displaying selected record details, and (c) successful CSV export from the history page.

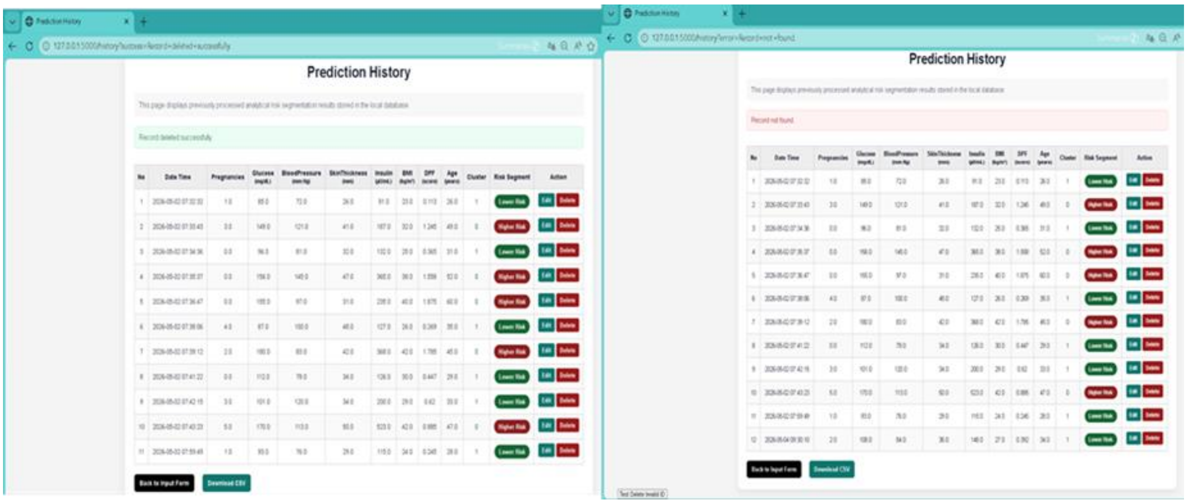


Figure 9b. Representative Black-Box Functional Testing Evidence: (d) confirmed deletion with success message and (e) non-existent-record handling.

DISCUSSION

Meaning of the Identified Risk Segments

The selected two-cluster structure produced an interpretable form of analytical risk segmentation. The Higher Risk segment showed higher mean Glucose, BMI, and Age, together with a higher proportion of Outcome = 1, whereas the Lower Risk segment showed lower values on those indicators. This pattern indicates that the final configuration separated the dataset into two distinct health-related profiles that were meaningful for monitoring. These groups should not be interpreted as diagnostic categories because they represent analytical segments derived from unsupervised clustering. Risk labels in this prototype were assigned based on the final cluster-label mapping derived from cluster profiling, not from single-variable clinical thresholds. Therefore, an individual input record may be assigned to the Lower Risk segment when its overall multivariate pattern is closer to the Lower Risk cluster centroid. This interpretation is consistent with previous diabetes-related clustering studies that identified subgroups with different metabolic and risk profiles (Onthoni et al., 2025; Taurbekova et al., 2025).

Contribution of the Web-Based Prototype for Monitoring

The main contribution of this study lies in integrating analytical risk segmentation into a web-based prototype. Many digital health systems still emphasize data entry, storage, and visualization, whereas the present prototype combined preprocessing, clustering, label mapping, storage, and presentation in one operational workflow. This made the system more than a recording tool because health-related input could be translated directly into an interpretable segmentation result. The result page, history module, CSV export, and dashboard supported structured monitoring and record visibility. The dashboard was also operationally relevant because it displayed current local prediction records stored in the local SQLite database rather than the raw research dataset. The history and dashboard pages strengthened traceability because each processed record retained its input values, unit labels, cluster ID, and badge-based risk segment. This contribution supports monitoring and interpretation, not clinical diagnosis (Kerr et al., 2024; Samadbeik et al., 2024).

Comparison with Previous Studies

This study bridges two research streams that are often developed separately. Previous web-based monitoring studies have mainly focused on data access, recording, and visualization, while clustering studies have mainly focused on subgroup identification and pattern discovery. The present study combined both by operationalizing K-Means-based analytical risk segmentation in a working web-based prototype. In this sense, its contribution is not only analytical but also system-oriented. The study also differs from prior Pima-based studies such as Chang et al. (2023) and Feng et al. (2023), which were oriented toward supervised classification. By contrast, the present study used unsupervised clustering to produce analytical risk segments and embedded those results in a monitoring prototype. It therefore offers a complementary contribution that is more appropriate for structured monitoring than for diagnostic classification.

CONCLUSION

This study successfully developed a web-based prototype for diabetes risk monitoring through analytical risk segmentation. The final K-Means configuration used the Extended Scenario at $k = 2$ and produced two interpretable analytical risk segments. Cluster 0 was interpreted as Higher Risk and contained 270 records, or 35.16% of the dataset, while Cluster 1 was interpreted as Lower Risk and contained 498 records, or 64.84% of the dataset. The Higher Risk segment showed higher mean Glucose, BMI, Age, and Outcome = 1 proportion than the Lower Risk segment. The prototype integrated preprocessing, K-Means cluster assignment, risk-label mapping, local SQLite storage, history management, CSV export, and dashboard visualization into one operational workflow. Manual black-box testing across 27 functional test cases showed that all tested features operated successfully. The main contribution of this study is the integration of clustering-based analytical segmentation into a working web-based monitoring prototype, rather than presenting clustering results only as a standalone data analysis output. The prototype supports monitoring and interpretation of diabetes-related risk patterns, but it does not provide clinical diagnosis. Future research should validate the approach using local clinical data, involve healthcare experts in risk interpretation, and conduct field-based usability testing with real users.

LIMITATIONS

This study has several limitations that affect the scope and interpretation of its findings. First, the analysis used a public secondary dataset rather than local clinical data. This limits the contextual relevance of the resulting segments to real local healthcare settings. Second, the dataset did not provide clinically validated risk labels for supervised diagnosis. As a result, the clustering output represents analytical risk segmentation rather than clinical diagnosis. Third, the system remains a web-based prototype rather than a deployed clinical platform. Its practical contribution should therefore be understood at the prototype level. Fourth, system testing focused on manual black-box functional testing, not clinical validation or user-centered field deployment. This means the study confirmed functional performance of the implemented routes and features, but not clinical usefulness in real practice. Fifth, the dashboard reflects stored prototype prediction records in the local SQLite database rather than the raw research dataset or a live clinical data source. This bounds the operational interpretation of the dashboard to prototype use. Sixth, the prototype did not implement expert-defined clinical threshold validation for each attribute; therefore, the input validation should be interpreted as a technical control for acceptable numeric ranges. These limitations define the actual contribution of the study and indicate that broader data, stronger validation, and more advanced implementation are still needed in future work.

REFERENCES

Abousaber, I., Abdallah, H. F., & El-Ghaish, H. (2025). Robust predictive framework for diabetes classification using optimized machine learning on imbalanced datasets. *Frontiers in Artificial Intelligence*, *7*, 1499530. https://doi.org/10.3389/frai.2024.1499530

Chang, V., Bailey, J., Xu, Q. A., & Sun, Z. (2023). Pima Indians diabetes mellitus classification based on machine learning algorithms. *Neural Computing and Applications*, *35*(22), 16157–16173. https://doi.org/10.1007/s00521-022-07049-z

Chicco, D., Campagner, A., Spagnolo, A., Ciucci, D., & Jurman, G. (2025). The Silhouette coefficient and the Davies-Bouldin index are more informative than Dunn index, Calinski-Harabasz index, Shannon entropy, and Gap statistic for unsupervised clustering internal evaluation of two convex clusters. *PeerJ Computer Science*, *11*, e3309. https://doi.org/10.7717/peerj-cs.3309

scikit-learn developers. (2026a). calinski_harabasz_score. *Scikit-learn*. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.calinski_harabasz_score.html

scikit-learn developers. (2026b). davies_bouldin_score. *Scikit-learn*. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.davies_bouldin_score.html

scikit-learn developers. (2026c). KMeans. *Scikit-learn*. https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html

scikit-learn developers. (2026d). MinMaxScaler. *Scikit-learn*. https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html

scikit-learn developers. (2026e). Model persistence. *Scikit-learn*. https://scikit-learn.org/stable/model_persistence.html

scikit-learn developers. (2026f). silhouette_score. *Scikit-learn*. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_score.html

scikit-learn developers. (2026g). SimpleImputer. *Scikit-learn*. https://scikit-learn.org/stable/modules/generated/sklearn.impute.SimpleImputer.html

International Diabetes Federation. (2025). *IDF Diabetes Atlas* (11th ed.). https://diabetesatlas.org/resources/idf-diabetes-atlas-2025/

Feng, X., Cai, Y., & Xin, R. (2023). Optimizing diabetes classification with a machine learning-based framework. *BMC Bioinformatics*, *24*, 428. https://doi.org/10.1186/s12859-023-05467-x

Ikotun, A. M., Habyarimana, F., & Ezugwu, A. E. (2025). Benchmarking validity indices for evolutionary K-means clustering performance. *Scientific Reports*, *15*, 21842. https://doi.org/10.1038/s41598-025-08473-6

Badan Kebijakan Pembangunan Kesehatan. (2024). *Prevalence, impact, and efforts in controlling hypertension and diabetes in Indonesia [Fact sheet]*. https://repository.badankebijakan.kemkes.go.id/id/eprint/5531/1/07%20factsheet%20PTM_English.pdf

Kerr, D., Ahn, D., Waki, K., Wang, J., Breznen, B., & Klonoff, D. C. (2024). Digital interventions for self-management of type 2 diabetes mellitus: Systematic literature review and meta-analysis. *Journal of Medical Internet Research*, *26*, e55757. https://doi.org/10.2196/55757

Onthoni, D. D., Chen, Y. E., Lai, Y. H., Li, G. H., Zhuang, Y. S., & Lin, H. M. (2025). Clustering-based risk stratification of prediabetes populations: Insights from the Taiwan and UK Biobanks. *Journal of Diabetes Investigation*, *16*(1), 25–35. https://doi.org/10.1111/jdi.14328

World Health Organization. (2024). *Urgent action needed as global diabetes cases increase four-fold over past decades*. https://www.who.int/news/item/13-11-2024-urgent-action-needed-as-global-diabetes-cases-increase-four-fold-over-past-decades

Plotnikova, V., Dumas, M., & Milani, F. P. (2022). Applying the CRISP-DM data mining process in the financial services industry: Elicitation of adaptation requirements. *Data and Knowledge Engineering*, *139*, 102013. https://doi.org/10.1016/j.datak.2022.102013

Samadbeik, M., Engstrom, T., Lobo, E. H., Kostner, K., Austin, J. A., Pole, J. D., & Sullivan, C. (2024). Healthcare dashboard technologies and data visualization for lipid management: A scoping review. *BMC Medical Informatics and Decision Making*, *24*(1), 352. https://doi.org/10.1186/s12911-024-02730-w

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, *181*, 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Tan, S. Y., Sumner, J., & Wang, Y. (2024). A systematic review of the impacts of remote patient monitoring interventions. *Npj Digital Medicine*, *7*, 192. https://doi.org/10.1038/s41746-024-01182-w

Taurbekova, B., Sarsenov, R., Yaqoob, M. M., Atageldiyeva, K., Semenova, Y., Fazli, S., & Sarria-Santamera, A. (2025). Cluster analysis in diabetes research: A systematic review enhanced by a cross-sectional study. *Journal of Clinical Medicine*, *14*(10), 3588. https://doi.org/10.3390/jcm14103588

Zhou, B., Rayner, A. W., Gregg, E. W., Sheffer, K. E., Carrillo-Larco, R. M., & Bennett, J. E. (2024). Worldwide trends in diabetes prevalence and treatment from 1990 to 2022. *The Lancet*, *404*(10467), 2077–2093. https://doi.org/10.1016/S0140-6736(24)02317-1