

Paper Type: Original Article

An Integrated Mobile Health System for Medication Management and Molecular Drug-Drug Interaction Screening Using Transformer-Based Learning

Mohamed Wael ¹ , Amal Elsayed ¹ , Basma Fawzy ¹ , Mohamed Mansour ¹ , Mahmoud Said ¹  and Abdelrhman Fathy ^{1,*} 

¹ Faculty of Computers and Informatics, Zagazig University, Zagazig 44519, Sharqiyah, Egypt;

Emails: waelmahdym@gmail.com, aamalelsayed2010@gmail.com, basmafawzy004@gmail.com, mohamedmns2005@gmail.com, mahmoudsaid1704@gmail.com, asawmaa@fci.zu.edu.eg.

Received: 05 May 2026**Revised:** 19 Jul 2026**Accepted:** 02 Oct 2026**Published:** 06 Oct 2026

Abstract

Medication-management systems often separate reminder functions from interaction-risk screening, caregiver support, and evidence-bounded medication information. This study presents Awan, a mobile-health research prototype that integrates medication scheduling, dose logging, optical medication capture, caregiver monitoring, pharmacy lookup, nutrition support, and drug-drug interaction (DDI) screening within a Flutter-Django client-server architecture. The DDI module compares name-based classical baselines, Morgan-fingerprint XGBoost models, and a ChemBERTa transformer operating on SMILES representations. On the recorded stratified pair-level test partition ($n = 65,621$), ChemBERTa produced 19,238 true negatives, 2,029 false positives, 3,336 false negatives, and 41,018 true positives, yielding accuracy 0.9182, positive-class precision 0.9529, recall 0.9248, F1-score 0.9386, specificity 0.9046, balanced accuracy 0.9147, and ROC-AUC 0.9635. Because class-0 examples were sampled from pairs not annotated as positive interactions, they are treated as operational pseudo-negatives rather than clinically proven non-interactions; because splitting occurred at the drug-pair level, the experiment establishes held-out pair performance rather than unseen-drug generalization. The scientific contribution is a safety-aware system integration that couples adherence and caregiver workflows with an experimentally evaluated molecular DDI service and an evidence-bounded herbal information layer while explicitly separating informational support from prescribing or therapeutic substitution. The prototype demonstrates the feasibility of this architecture and defines the validation boundaries required for responsible clinical translation.

Keywords: Medication Adherence; Mobile Health; Drug-drug Interaction; Chemberta; Molecular Representation; Caregiver Monitoring; Clinical Decision Support; Digital Health.

1 | Introduction

Long-term medication use places a sustained cognitive and organizational burden on patients, particularly when treatment involves multiple medicines, different dosing intervals, changing prescriptions, and communication with family members or caregivers. The World Health Organization identified poor



Corresponding Author: asawmaa@fci.zu.edu.eg



<https://doi.org/10.61356/j.scin.2026.3674>



Licensee SciNexuses. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0>).

adherence to long-term therapy as a major global health problem and reported that, in developed settings, adherence among patients with chronic disease averaged approximately 50%, with lower levels reported in some developing settings [1]. More recent systematic reviews continue to show that well-designed mobile interventions can improve medication-adherence outcomes, although effects depend on population, intervention design, adherence measurement, and study quality [2-4].

Medication management is not only a reminder problem. A patient may take a dose at the wrong time, combine interacting products, confuse trade names, fail to recognize an adverse-effect warning, or use an herbal product together with prescribed medication. Digital support must therefore separate simple scheduling from higher-risk decision-support functions. This distinction is especially important for herbal products, because the evidence base for many herb–drug interactions remains incomplete and heterogeneous; authoritative guidance explicitly cautions that some interactions are supported only indirectly while others are clinically important [12].

1.1 | Problem Definition and Research Objective

The objective of Awan is to implement a unified mobile platform in which medication organization, dose reminders, adherence records, caregiver support, medication-information retrieval, and DDI screening can be accessed through one patient-facing workflow. The project also explores how molecular machine learning can be incorporated into a mobile-health backend without presenting the model as a replacement for professional clinical judgment.

1.2 | Research Contributions

The scientific contribution is system-level rather than a claim of a new molecular foundation model. First, Awan defines an end-to-end mobile-health architecture in which medication schedules, dose-event logging, patient-authorized caregiver links, and DDI services operate over a shared medication state instead of as disconnected utilities. Second, it evaluates three representation families for DDI screening—categorical drug-name baselines, Morgan fingerprints with XGBoost, and transformer-based SMILES representation—within the same application context. Third, it makes label semantics and evaluation scope explicit by distinguishing pseudo-negative pairs from clinically verified non-interactions and by limiting pair-split conclusions to held-out combinations rather than novel drugs. Fourth, it embeds higher-risk functions, including DDI warnings, herbal information, nutrition guidance, and conversational assistance, inside an explicit human-in-the-loop safety boundary. Together, these elements make the work a translational systems contribution linking mHealth workflow design, molecular machine learning, and responsible medication-safety communication.

2 | Related Work and Design Positioning

2.1 | Mobile Medication-Adherence Interventions

Evidence from randomized trials suggests that mobile applications can support adherence through combinations of reminders, documentation, feedback, education, data sharing, and clinical decision-support features. Peng et al. reported a significant pooled improvement in medication adherence across trials of adults with chronic disease, while a later systematic review by Lanke et al. found improvement across multiple chronic conditions but also emphasized heterogeneity in adherence measures and intervention designs [2], [3]. An umbrella review covering 34 meta-analyses likewise found meaningful benefits for selected chronic-disease outcomes while noting that many pooled effects were not significant [4]. These findings support the use of mobile tools as an adjunct to care, but they do not justify assuming that any specific prototype improves clinical outcomes without prospective evaluation.

2.2 | Existing Digital Medication Services

Commercial and regional services demonstrate different strengths. Medisafe publicly documents medication scheduling, adherence tracking, drug-interaction alerts, progress reporting, and caregiver support through Medfriend or dependent-profile workflows [5]. Yodawy focuses on prescription access, chronic medication refills, pharmacy fulfillment, and insurance-linked services in Egypt [6]. Chefaa combines pharmacy delivery, prescription upload, medication reminders or refill support, and pharmacist communication [7]. Awan is positioned as a research prototype rather than as a claim of commercial superiority: its distinguishing research scope is the integration of adherence and caregiver workflows with an experimentally evaluated molecular DDI service and an evidence-bounded herbal information layer. Table 1 summarizes this functional positioning using only features documented by the cited public sources.

Table 1. Functional positioning of Awan relative to representative medication-management services.

System / service	Primary documented focus	Caregiver / adherence support	DDI or herb-focused research component
Medisafe	Medication scheduling, adherence/progress tracking, measurements, and interaction alerts [5]	Medfriend/dependent workflows and missed-dose notifications are documented [5]	Public documentation describes interaction alerts, but does not document the molecular-model evaluation or evidence-bounded herbal layer studied here.
Yodawy	Prescription access, monthly chronic refills, pharmacy fulfillment, and insurance-linked services in Egypt [6]	Supports continuity of chronic medication supply; the cited public page emphasizes access and fulfillment [6]	The cited public page does not document a molecular DDI research pipeline or an integrated herb-safety layer.
Chefaa	Pharmacy delivery, prescription upload, treatment/refill reminders, and pharmacist-facing support [7]	Reminder/refill support is documented; a patient-caregiver monitoring dashboard is not documented on the cited public pages	The cited service description focuses on access and reminders rather than molecular DDI modeling.
Awan	Medication scheduling and logging, scanning, pharmacy lookup, caregiver monitoring, DDI screening, and evidence-linked medication/herbal information	Patient-authorized caregiver connection, adherence summaries, and irregularity alerts are integrated in the prototype	Includes a molecular DDI evaluation branch (fingerprints and ChemBERTa) and an evidence-bounded herbal information layer.

Table 1 shows that the contribution does not rest on a single consumer feature such as reminders, pharmacy delivery, caregiver alerts, or interaction warnings, all of which exist in different forms elsewhere. The research distinction is the coordinated use of a shared medication state across adherence logging, caregiver escalation, molecular DDI screening, and evidence-bounded medication/herbal information. The comparison is functional and descriptive; it is not a clinical-effectiveness ranking.

2.3 | Molecular DDI Prediction and Evaluation

Computational DDI prediction has progressed from hand-crafted descriptors and molecular fingerprints toward deep neural, graph-based, knowledge-graph, multimodal, and self-supervised representations [29], [30]. ChemBERTa is relevant to the present work because it applies large-scale transformer pretraining to SMILES strings and provides a learned alternative to fixed molecular fingerprints [17]. This broader literature also makes clear that high benchmark scores alone are insufficient for a safety-oriented interpretation: label quality, negative-sample construction, calibration, external validation, and the definition of the test distribution directly determine what a performance estimate means [29, 30].

Evaluation protocol is particularly important for DDI models. Kpanou et al. distinguish random pair splits from one-unseen-drug and both-unseen-drug settings and show that performance can decline materially as

evaluation moves toward novel-drug scenarios [28]. A recent systematic review likewise identifies heterogeneous and sometimes optimistic split protocols, limited external validation, and uncertainty about translation as persistent issues in DDI research [30]. These findings motivate the conservative validity scope used in this paper: the reported Awan experiment is interpreted as pair-level discrimination on its stated split, not as evidence of cold-start performance for unseen compounds.

3 | System Architecture and Functional Design

3.1 | Client–Server Architecture

Awan follows a client-server architecture. The mobile client is implemented in Flutter/Dart, while the backend uses Django REST Framework. The mobile layer handles user interaction, on-device scheduling, local storage, camera capture, and presentation of adherence information. The backend manages authentication, medication records, adherence events, caregiver links, API-level business logic, database operations, and calls to AI or reference-data services. PostgreSQL is used as the server-side relational store and SQLite/SharedPreferences are used for local device storage. JSON over HTTPS is the intended API exchange format, and JSON Web Tokens are used for authenticated sessions. Figure 1 summarizes the separation between the patient/caregiver interfaces, the application backend, persistence, notification services, and the DDI intelligence layer.

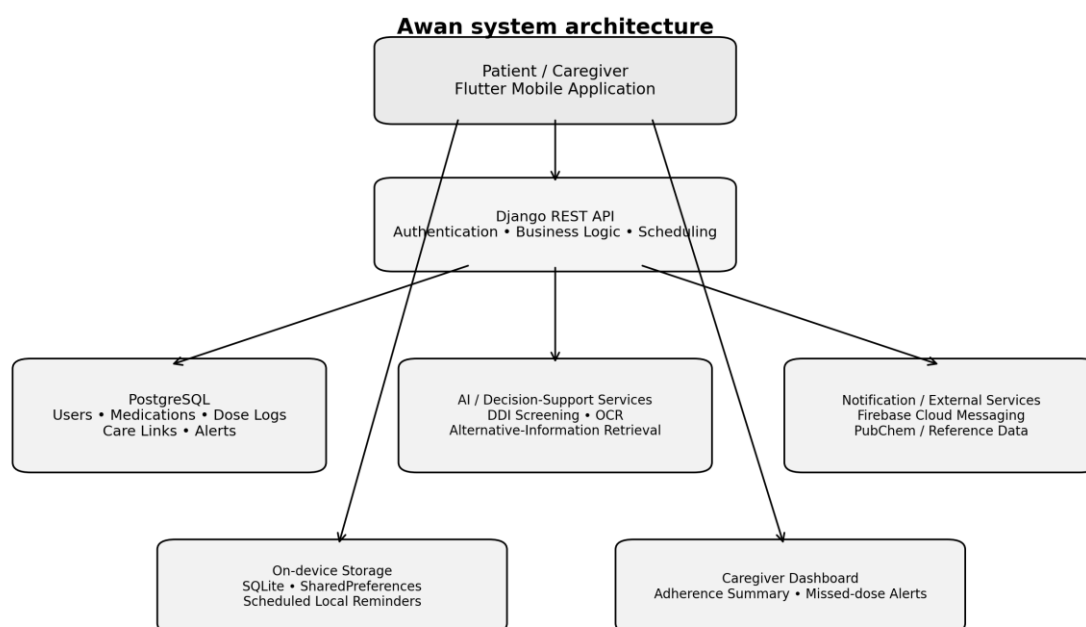


Figure 1. Layered architecture of the Awan application.

3.2 | Core Functional Modules

The medication-management module stores medication name, dosage, timing, frequency, and treatment duration. The reminder module schedules local notifications and records whether a scheduled dose was marked as taken, missed, or unresolved. The camera/OCR workflow captures a medication label to accelerate information retrieval. The DDI module evaluates a medication pair and returns a model score together with a warning-oriented display. The caregiver module links a care person to a patient-controlled account and exposes adherence summaries and irregularity alerts. Additional modules provide nearby-pharmacy lookup, medication information, an informational herbal layer, a nutrition assistant, and a general health-information chatbot. Figure 2 traces how these branches originate from the authenticated user record and converge back

to a common dashboard, which is important because the modules share the same medication and adherence state rather than operating as independent tools.

3.3 | Adherence Representation

For a patient u over an observation window T , the software-level adherence ratio is defined as the proportion of scheduled doses explicitly logged as taken. This operational measure is useful for in-app monitoring but should not be treated as a validated clinical adherence instrument unless compared against an accepted reference method.

$$A_u(T) = \frac{N_{\text{taken}}(u,T)}{N_{\text{scheduled}}(u,T)} \quad (1)$$

A missed-dose alert can be generated when a scheduled event remains unconfirmed beyond an application-defined grace interval. The caregiver view should display the underlying event status rather than infer nonadherence from a missing log alone, because a missing digital confirmation can also reflect device or connectivity behavior.

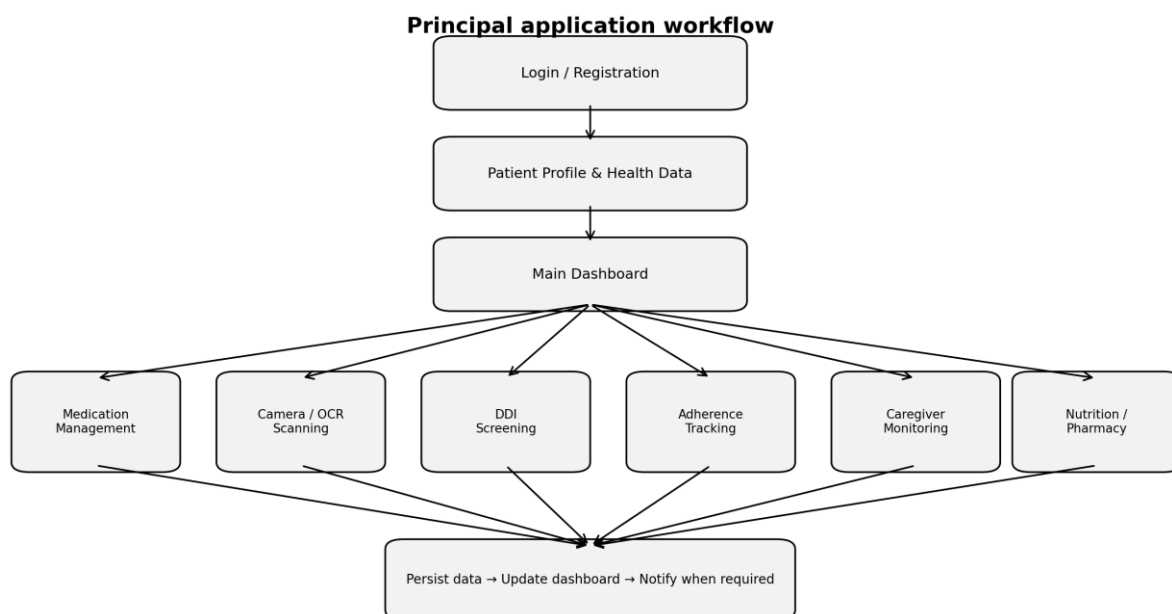


Figure 2. Principal application workflow and functional branches.

3.4 | Security and Access-Control Considerations

The prototype uses authenticated API access and separates patient and caregiver roles. Caregiver monitoring is meaningful only when linking is patient-authorized, revocable, and restricted to the minimum information needed for support. Health records, medication histories, care links, model outputs, and alert events are sensitive data; production deployment therefore requires secure transport, least-privilege authorization, credential protection, server hardening, audit logging, backup and recovery policies, privacy notices, and jurisdiction-specific regulatory review. These are deployment requirements, not claims of certification in the present prototype.

4 | Data and Methods

4.1 | Data Sources

The project combines data sources with distinct operational roles. DDI supervision is derived from the Mendeley Data resource extracted from DrugBank 5.1 [8], [9]. Molecular structures are obtained from DrugBank records and, when necessary, resolved through PubChem [10]. Medicine-substitute and active-ingredient tables were obtained from public Kaggle datasets [26], [27]. Active ingredients are mapped to Anatomical Therapeutic Chemical (ATC) classes using the WHO ATC/DDD system [11]. Herbal candidate entries originate from the project's traditional-medicine source [14], but safety-facing information is cross-checked against contemporary sources such as MedlinePlus and NCCIH [12], [13]. A local pharmacy table for Zagazig was assembled from a public business directory and map-derived location information [15], [16]. Table 2 separates the role of each source from the validity boundary attached to that source.

Table 2. Data sources, system roles, and interpretation limits.

Dataset / source	Primary role in Awan	Key fields used	Validity boundary
DrugBank-derived DDI dataset [8], [9]	Positive DDI supervision	Drug identifiers, names, interaction type	Known interaction entries are positives; absence from the source is not proof of no interaction.
DrugBank / PubChem [9], [10]	Molecular structure resolution	SMILES / chemical identifiers	Structure resolution errors must be filtered before modeling.
Medicine datasets [26], [27]	Trade-name, ingredient and substitute lookup	Drug name, active ingredient, substitute	Commercial availability and equivalence require local clinical/pharmacy verification.
WHO ATC/DDD [11]	Therapeutic-class normalization	ATC code and level-1 class	ATC class similarity does not establish therapeutic interchangeability.
MedlinePlus/NCCIH + traditional compendium [12]-[14]	Herbal candidate generation, evidence notes, and safety annotation	Herb name, claimed use, warning, evidence note	Traditional material is not treated as a safety authority; herbal entries are informational and require contemporary evidence review.
Zagazig pharmacy directory + maps [15], [16]	Local pharmacy lookup	Name, address, coordinates, contacts	Coverage is local and time-sensitive; listings can change.

The source heterogeneity in Table 2 is deliberate but requires source-specific interpretation. Curated DDI records can support supervised interaction labels; ATC codes support classification but not interchangeability; medicine-market datasets support lookup but not therapeutic equivalence; and herbal sources support evidence-linked education rather than autonomous substitution. This separation prevents a database join from being mistaken for clinical validation.

4.2 | DDI Label Construction and Preprocessing

The source DDI table primarily supplies known interacting pairs. The pipeline generates additional candidate pairs by randomly pairing drugs and excluding pairs already present in the positive interaction set. These generated pairs are operational pseudo-negatives or unlabeled pairs, not clinically verified non-interactions, because an absent database edge can reflect incomplete knowledge rather than true safety. This limitation is well recognized in DDI modeling, where unlabeled pairs may contain undiscovered positives [29], [30]. Figure 3 shows the resulting sequence from source harmonization and pseudo-negative construction through molecular representation, model training, and pair-level evaluation.

Duplicate drug pairs were removed and labels were converted to numeric form. SMILES strings were extracted from DrugBank XML and unresolved structures were queried through PubChem. During intermediate data assembly, the literal string 'C' was used as a missing-structure sentinel; the retained preprocessing logic removes rows in which either molecular field equals 'C' before model training. Accordingly, the sentinel is treated as missing data and is never interpreted as a valid methane representation in the learning set.

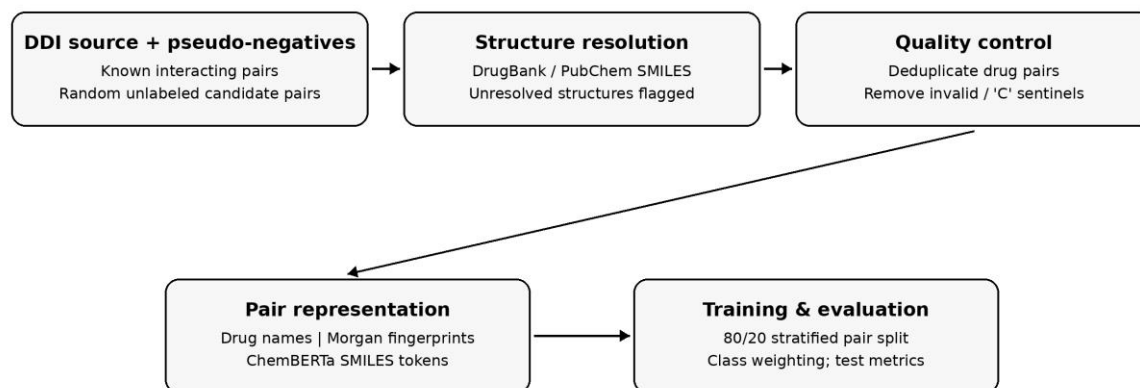


Figure 3. DDI data preparation, representation, and evaluation pipeline.

4.3 | Model Families

4.3.1 | Drug-Name Baselines

Classical baseline models represent the two drug names as categorical variables and use standard classifiers such as logistic regression, random forest, or XGBoost. This baseline is computationally inexpensive but can memorize identities and co-occurrence patterns without encoding molecular structure. Its apparent performance under random pair splitting can therefore be optimistic.

4.3.2 | Morgan-Fingerprint XGBoost

For structure-aware classical learning, each SMILES string is converted to an extended-connectivity (Morgan) fingerprint using RDKit [21-25]. Two drug fingerprints are concatenated to form the pair representation, which is classified using XGBoost [18, 19]. The tuned configuration reported in the project uses 500 estimators, maximum tree depth 9, learning rate 0.03, and subsample ratio 0.9.

4.3.3 | ChemBERTa Transformer

ChemBERTa applies transformer-based representation learning to tokenized SMILES strings [17]. In the project, paired molecular inputs are tokenized with a maximum sequence length of 128 and passed through the pretrained transformer followed by a classification head. The recorded training configuration uses a learning rate of 2×10^{-5} , batch size 32, five epochs, weight decay 0.01, and AdamW optimization [20]. The transformer branch is therefore evaluated as a learned molecular representation strategy rather than as a newly proposed encoder architecture.

4.4 | Class Weighting and Objective Function

Because the operational binary dataset is imbalanced, the transformer uses weighted cross-entropy. For N training pairs, true class y_i , predicted class probability $p(y_i | x_i)$, and class weight w_{y_i} , the optimization objective is:

$$L = -\frac{1}{N} \sum_{i=1}^N w_{y_i} \log p(y_i | x_i) \quad (2)$$

SMOTE/undersampling was explored for classical models, whereas weighted loss was used for ChemBERTa. The use of class weighting changes the training objective but does not eliminate the need to report class-specific metrics.

4.5 | Train/Test Protocol and Validity Scope

The reported experiments use an 80/20 stratified split at the drug-pair level. Stratification preserves the operational class ratio, but pair-level splitting can place the same individual drug in both training and test partitions. Consequently, the results evaluate generalization to held-out pairs under this split; they do not establish performance for one-unseen-drug, both-unseen-drug, scaffold-disjoint, temporal, or prospective clinical settings. This distinction follows established DDI evaluation analyses showing that random pair performance and cold-start drug generalization are materially different questions [28], [30].

4.6 | Evaluation Metrics

Performance is summarized using accuracy, precision, recall, F1-score, ROC-AUC, and the confusion matrix. With TP, TN, FP, and FN denoting the usual binary outcomes, the principal point metrics are:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (5)$$

5 | Results

5.1 | Model-Family Comparison

Table 3 summarizes the model-family results recorded under the project's stratified pair-level evaluation. The exploratory baseline and XGBoost values were retained as rounded aggregate figures, whereas the ChemBERTa metrics can be verified directly from the retained confusion matrix and classification output. The table is therefore used for descriptive comparison of representation strategies rather than for a formal statistical ranking across models.

Table 3. Model-family results under the reported stratified pair-level split.

Model	Input representation	Accuracy	F1-score	ROC-AUC	Interpretation
Name-based baseline	Drug names / categorical	≈0.97	≈0.98	Not reported	High aggregate score, but the recorded analysis reported weak class-0 recall; identity memorization remains a concern.
Fingerprint XGBoost	Morgan fingerprints	≈0.90	≈0.90	Not reported	Introduces molecular structure through a fixed handcrafted representation.
Tuned XGBoost	Morgan fingerprints	≈0.954	≈0.954	≈0.989	Strong reported pair-split classical performance; exact repeated-run uncertainty is not available in the retained summary.
ChemBERTa	Raw/tokenized SMILES	0.9182	0.9386	0.9635	Exact metrics are verifiable from the recorded confusion matrix; learned SMILES representation avoids manual fingerprint construction.

5.2 | ChemBERTa Confusion-Matrix Analysis

As shown in Figure 4 and Table 4, the ChemBERTa test output contains 65,621 pairs: 21,267 operational class-0 pairs and 44,354 class-1 known-interaction pairs. The confusion matrix contains 19,238 true negatives, 2,029 false positives, 3,336 false negatives, and 41,018 true positives. These counts yield accuracy 0.9182, positive-class precision 0.9529, recall 0.9248, F1-score 0.9386, specificity 0.9046, and balanced accuracy 0.9147; the Matthews correlation coefficient is 0.8171. The recorded ROC-AUC is 0.9635. Figure 5 summarizes the principal scalar metrics. A Wilson 95% confidence interval for the observed pair-level accuracy is 0.9161-0.9203; this interval reflects sampling uncertainty for the held-out pair partition only and is not a clinical-performance interval. The corresponding false-negative rate is 7.52% and the false-positive rate is 9.54%, both of which are operationally relevant in a safety-screening context.

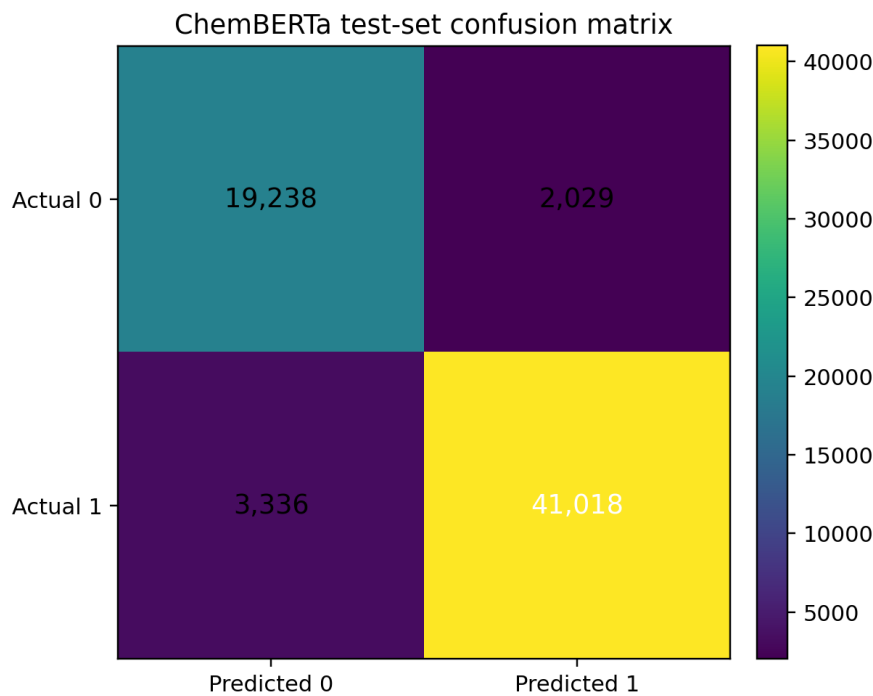


Figure 4. ChemBERTa confusion matrix on the pair-level test partition.

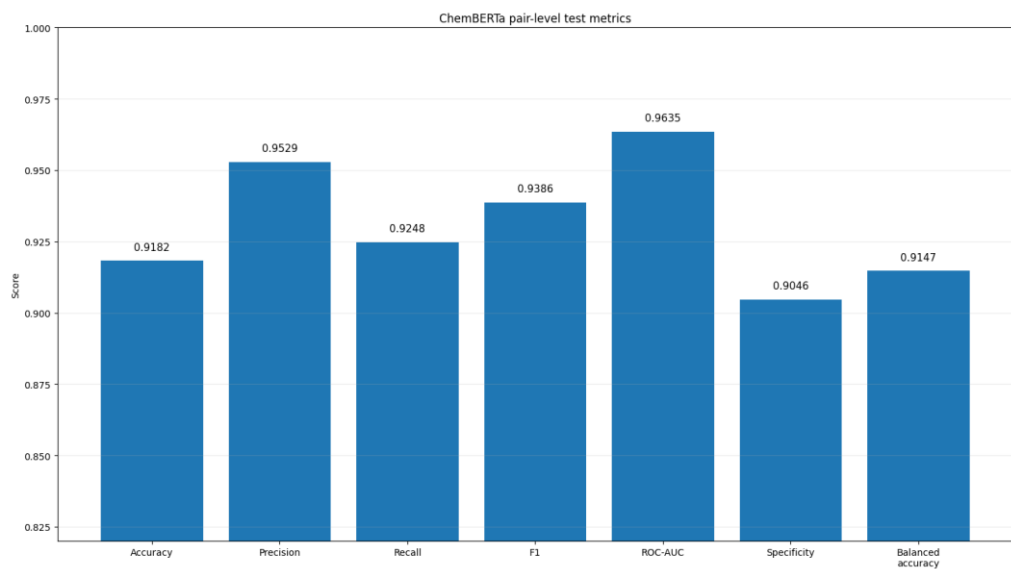


Figure 5. ChemBERTa pair-level test metrics.

Table 4. ChemBERTa class-specific metrics on the pair-level test partition.

Class / aggregate	Precision	Recall	F1-score	Support
Class 0 (operational pseudo-negative)	0.85	0.90	0.88	21,267
Class 1 (known interaction)	0.95	0.92	0.94	44,354
Macro average	0.90	0.91	0.91	65,621
Weighted average	0.92	0.92	0.92	65,621

5.3 | Interpretation of the Model Results

The available pair-level summary does not justify labeling ChemBERTa as numerically superior to the tuned fingerprint model: the tuned XGBoost row reports higher aggregate accuracy and ROC-AUC. ChemBERTa nevertheless offers a different engineering trade-off by learning a SMILES representation end to end rather than depending on manually constructed fingerprints. The application architecture is model-agnostic, so model selection should be based on repeated seed-controlled evaluation, calibration, computational cost, and the intended deployment distribution rather than on model family alone.

The false-negative count is especially important for a safety-oriented DDI screen. A false negative corresponds to a known interaction in the operational label set being predicted as non-interacting. Threshold selection should therefore be tuned to the intended screening role, and a deployed service should combine model output with deterministic interaction knowledge and professional review rather than use the classifier as a standalone authorization mechanism.

5.4 | Functional Prototype

The implemented interfaces demonstrate the end-to-end user workflow: medication entry and detail review, camera-assisted label capture, pharmaceutical and herbal information search, personalized nutrition output, nearby-pharmacy lookup, and caregiver linking. Figure 6 presents representative screens from these implemented modules. The interfaces provide evidence that the system components are integrated at the prototype level; they are not presented as evidence of clinical effectiveness or human-factors validation.

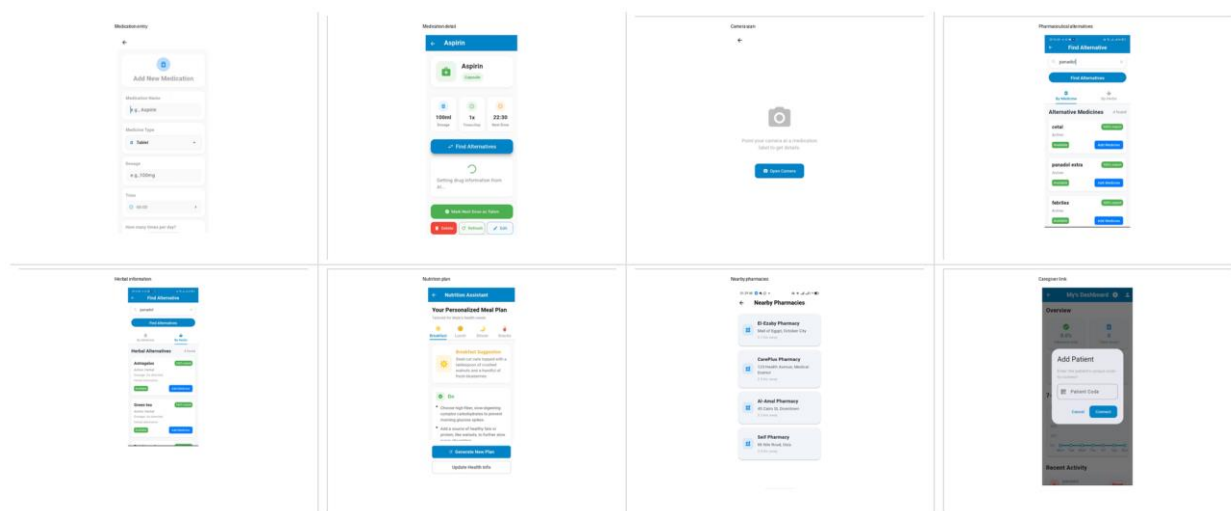


Figure 6. Representative Awan mobile interfaces: medication management, scanning, medication/herbal information, nutrition, pharmacy lookup, and caregiver linking.

5.5 | Scientific Contribution and System-Level Significance

Awan's main scientific value lies in the coupling between system workflow and model governance. A single medication state drives scheduled reminders, dose-event logging, caregiver escalation, interaction screening, and evidence-linked information retrieval. This architecture allows high-risk outputs to inherit the same patient context used by adherence functions while preserving an explicit boundary between convenience

features and clinical decision support. In practical terms, a missed-dose event, a newly added medication, and a DDI warning are not isolated notifications; they are coordinated events within one auditable medication-management workflow.

This positioning differs from two common research patterns. Compared with reminder- or fulfillment-centered mobile services, Awan exposes the assumptions and error profile of the molecular DDI layer rather than treating interaction checking as a black box. Compared with stand-alone DDI modeling studies, it places the prediction service inside a patient/caregiver workflow with explicit evidence, escalation, and non-substitution rules. The contribution is therefore translational systems design—integrating molecular prediction with medication-management operations and safety semantics—rather than a claim of a new transformer architecture or proven superiority in clinical outcomes.

6 | Clinical Safety, Reliability, and Responsible Use

6.1 | Decision-Support Boundary

Awan is a software decision-support prototype. It does not establish a diagnosis, authorize medication substitution, change a prescribed dose, or replace a physician or pharmacist. DDI output should be displayed as a screening signal linked to its evidence source, model confidence or decision threshold, and an appropriate action message. High-risk warnings should direct the user toward professional review rather than instruct the user to stop, replace, or initiate treatment autonomously.

6.2 | Herbal Information Boundary

The herbal module requires especially careful wording. An herb mapped to a therapeutic class is not pharmacologically equivalent to a prescription medicine, and an interface 'match' value cannot establish efficacy, safety, dose equivalence, or interchangeability. NCCIH notes that evidence for many herb-drug interactions remains limited or indirect, while some interactions are clinically important [12]. The system therefore treats herbal content as evidence-linked educational information with explicit warnings and source provenance, never as an autonomous substitution recommendation.

6.3 | Nutrition and Conversational AI

Nutrition and conversational outputs are likewise non-diagnostic. Any diet recommendation that depends on disease state, renal or hepatic function, pregnancy, allergy, or drug-food interactions requires validated clinical rules and qualified review. Conversational assistance should prioritize education, navigation of the user's medication record, and escalation when a medication question is uncertain or potentially urgent; it should not generate unsupervised prescribing, dose changes, or emergency triage conclusions.

6.4 | Functional Reliability

Reliability has two distinct meanings in this project. Software reliability concerns API behavior, persistence, notification scheduling, role authorization, data synchronization, and error handling; model reliability concerns discrimination and calibration on a defined evaluation distribution. A high ROC-AUC does not guarantee reliable application operation, and successful API tests do not establish clinical model validity. Treating these dimensions separately prevents software correctness from being mistaken for clinical evidence.

7 | Limitations and Threats to Validity

The DDI experiment has several limitations that define, rather than invalidate, its scope. First, class-0 examples are sampled from pairs absent from the known-positive set, so they are pseudo-negatives rather than clinically confirmed non-interactions. Second, the 80/20 pair-level split can share individual drugs across partitions and can overstate performance for novel-drug scenarios [28], [30]. Third, the reported model-family comparison is based on a single project evaluation, with rounded baseline summaries rather than repeated-

seed distributions. Fourth, labels are database-derived and do not represent prospectively adjudicated clinical outcomes. These constraints mean the reported metrics should be interpreted as proof-of-concept pair-level model performance within the stated data construction.

At the application level, the study demonstrates functional integration rather than clinical effectiveness. It does not claim that Awan increases real-world adherence, reduces hospitalization, or prevents adverse events without prospective patient evaluation. The Zagazig pharmacy layer is geographically local and time-sensitive; the herbal layer requires continuous evidence review; and privacy, cybersecurity, regulatory classification, accessibility, and human-factors performance require formal assessment before clinical deployment. These boundaries are explicit design constraints for the next validation stage rather than unstated assumptions.

8 | Conclusion and Future Work

Awan integrates medication scheduling, adherence logging, caregiver support, medication scanning, pharmacy lookup, DDI screening, and evidence-bounded health-information services within a single mobile architecture. On the reported pair-level test partition, the ChemBERTa branch achieved accuracy 0.9182, precision 0.9529, recall 0.9248, F1-score 0.9386, specificity 0.9046, balanced accuracy 0.9147, and ROC-AUC 0.9635. More importantly, the study ties those model outputs to explicit label semantics, split validity, caregiver authorization, and non-substitution safety rules. The resulting contribution is a coherent mobile-health systems framework in which medication-management functions and molecular interaction screening share context without conflating software assistance with clinical authority.

The next validation stage should prioritize one-unseen-drug, both-unseen-drug, scaffold-disjoint, and repeated-seed evaluation; confidence intervals and calibration analysis; external DDI benchmarks; deterministic database verification alongside model scores; and threshold selection aligned with screening risk. Application validation should include usability and accessibility studies, security testing, prospective adherence measurement, clinician/pharmacist review of safety messages, and structured validation of herbal and nutrition content. These steps would test whether the technically integrated prototype can support reliable use in real clinical and caregiving workflows.

Author Contribution

All authors contributed equally to this study.

Funding

This research received no external funding.

Data Availability

Public data sources, identifiers, and their roles are listed in Table 2 and the reference list. Reuse of DrugBank-derived content remains subject to the applicable DrugBank access terms, while the Mendeley DDI dataset is distributed under its stated license [8]. The locally curated pharmacy layer is based on publicly listed business and map information and should be refreshed before reuse because locations and contact details can change. The exact ChemBERTa confusion counts and all derived point metrics used in this manuscript are reported in Section 5, enabling direct numerical audit of the test result. The reported study concerns software development and secondary/public drug and pharmacy information; it does not report prospective patient recruitment, a clinical intervention, or a patient-level clinical dataset.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] World Health Organization, *Adherence to Long-Term Therapies: Evidence for Action*. Geneva, Switzerland: WHO, 2003.
- [2] Y. Peng, H. Wang, Q. Fang, L. Xie, L. Shu, W. Sun, and Q. Liu, "Effectiveness of Mobile Applications on Medication Adherence in Adults with Chronic Diseases: A Systematic Review and Meta-Analysis," *J. Manag. Care Spec. Pharm.*, vol. 26, no. 4, pp. 550–561, 2020, doi: 10.18553/jmcp.2020.26.4.550.
- [3] V. Lanke, K. Trimm, B. Habib, and R. Tamblyn, "Evaluating the Effectiveness of Mobile Apps on Medication Adherence for Chronic Conditions: Systematic Review and Meta-Analysis," *J. Med. Internet Res.*, vol. 27, e60822, 2025, doi: 10.2196/60822.
- [4] S. Sun, O. Simonsson, S. McGarvey, J. Torous, and S. B. Goldberg, "Mobile phone interventions to improve health outcomes among patients with chronic diseases: an umbrella review and evidence synthesis from 34 meta-analyses," *Lancet Digit. Health*, vol. 6, no. 11, pp. e857–e870, 2024, doi: 10.1016/S2589-7500(24)00119-5.
- [5] Medisafe, "Download the App: Medication Management, Caregiver Support, and Drug Interaction Alerts," 2026. Accessed: Sep. 22, 2026. <https://medisafe.com/download-the-app>
- [6] Yodawy, "Healthcare Simplified, For Everyone," 2026. Accessed: Sep. 22, 2026. <https://yodawy.com/>
- [7] Chefaa, "About Chefaa: Digital Pharmacy, Medication Access, and Refill Support," 2026. Accessed: Sep. 22, 2026. <https://chefaa.com/eg-ar/about-us>
- [8] H. Yu, "data of multiple-type drug-drug interactions," Mendeley Data, Version 1, 2020, doi: 10.17632/md5czfsfnd.1.
- [9] D. S. Wishart et al., "DrugBank 5.0: a major update to the DrugBank database for 2018," *Nucleic Acids Res.*, vol. 46, no. D1, pp. D1074–D1082, 2018, doi: 10.1093/nar/gkx1037.
- [10] S. Kim et al., "PubChem in 2021: new data content and improved web interfaces," *Nucleic Acids Res.*, vol. 49, no. D1, pp. D1388–D1395, 2021, doi: 10.1093/nar/gkaa971.
- [11] World Health Organization, "Anatomical Therapeutic Chemical (ATC) classification system and Defined Daily Dose (DDD)," 2026. Accessed: Sep. 22, 2026.
- [12] National Center for Complementary and Integrative Health, "Herb-Drug Interactions," NIH, Dec. 2024. Accessed: Sep. 22, 2026.
- [13] U.S. National Library of Medicine, "Herbs and Supplements," MedlinePlus, National Institutes of Health. Accessed: Sep. 22, 2026.
- [14] D. Babli, Al-Tibb Al-Badil (Alternative Medicine), ch. 11. Riyadh, Saudi Arabia, 2007. Used in the project only as a traditional candidate-information source; safety statements are cross-checked against contemporary references.
- [15] Dalil Zagazig Business Directory, "Pharmacy Listings—Zagazig City," Egypt. Accessed: 2026.
- [16] Google LLC, "Google Maps." Accessed: 2026.
- [17] S. Chithrananda, G. Grand, and B. Ramsundar, "ChemBERTa: Large-Scale Self-Supervised Pretraining for Molecular Property Prediction," arXiv:2010.09885, 2020.
- [18] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
- [19] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.
- [20] I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization," in *Proc. ICLR*, 2019.
- [21] D. Rogers and M. Hahn, "Extended-Connectivity Fingerprints," *J. Chem. Inf. Model.*, vol. 50, no. 5, pp. 742–754, 2010.
- [22] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
- [23] A. Paszke et al., "PyTorch: An Imperative Style, High-Performance Deep Learning Library," in *Advances in Neural Information Processing Systems* 32, 2019.
- [24] G. Landrum, "RDKit: Open-Source Cheminformatics Software," 2006–2026. <https://www.rdkit.org/>
- [25] W. McKinney, "Data Structures for Statistical Computing in Python," in *Proc. 9th Python in Science Conf.*, 2010, pp. 56–61.
- [26] S. Singh, "250k Medicines Usage, Side Effects and Substitutes," Kaggle, 2023. Accessed: 2026.
- [27] S. Singh, "A–Z Medicine Dataset of India," Kaggle, 2023. Accessed: 2026.
- [28] R. Kpanou, M. A. Osseni, P. Tossou, F. Laviolette, and J. Corbeil, "On the robustness of generalization of drug-drug interaction models," *BMC Bioinformatics*, vol. 22, art. 477, 2021, doi: 10.1186/s12859-021-04398-9.
- [29] X. Li, Z. Xiong, W. Zhang, and S. Liu, "Deep learning for drug-drug interaction prediction: A comprehensive review," *Quantitative Biology*, vol. 12, no. 1, pp. 30–52, 2024, doi: 10.1002/qub.232.
- [30] Q. Hashir, M. A. E. Yar, A. Ullah, S. Kamal, F. Ullah, and Z. H. Abdullah, "Computational approaches for drug-drug interaction prediction: a systematic review of data sources, modeling strategies, and evaluation frameworks," *Front. Pharmacol.*, vol. 17, 2026, doi: 10.3389/fphar.2026.1816394.

Disclaimer/Publisher's Note: The perspectives, opinions, and data shared in all publications are the sole responsibility of the individual authors and contributors, and do not necessarily reflect the views of Sciences Force or the editorial team. Sciences Force and the editorial team disclaim any liability for potential harm to individuals or property resulting from the ideas, methods, instructions, or products referenced in the content.