UNIVERSITY OF THE PHILIPPINES
CENTER FOR INTEGRATIVE AND DEVELOPMENT STUDIES

POLICY BRIEF 2026-26
ISSN 2619-7278 (Print)
ISSN 2619-7286 (Online)

PROGRAM ON DATA SCIENCE FOR PUBLIC POLICY

AI-Enhanced Community Learning:

Using Machine Learning Algorithms to Redesign Republic Act No. 7743

Jonathan G. Medina[1], Brian Lloyd B. Dayrit[2], Maria Maura S. Tinao[3], and Corazon M. Gamboa[4]

Executive Summary

In 1994, Republic Act No. 7743 mandated the creation of a public library and reading center to encourage Filipinos to keep learning. To this date, for over 30 years, the system remains severely broken. Out of the registered libraries, only about 30 percent are open right now. The old library model does not work anymore because there are not enough resources, people do not know how to use technology, and there are not enough flexible technologies. This short guide shows how to use AI and machine learning to make RA No. 7743 work well. So, changing from a simple establishing mandate to a lively, performance-based, data-driven mandate will help the government close the digital divide, use its resources more effectively, and get ready for community learning in the future.

Problem Statement: The Stagnation of RA 7743

RA 7743 stipulated that there should be a library in every community, but this has not been the case because not all have equal infrastructure, budgets, and technology.

- High Rates of Inactivity: The 2025 National Library Directory presents over 1,500 public libraries; however, a considerable portion of these are either closed or simply not open.

- Regional Inequality: There is a big digital gap. The more urbanized areas, like NCR and Region I, have steady operations. BARMM, Caraga, and Region II have low operational capacities and are at risk of closure.

- Technological Lag: Nonavailability of IT services and internet access within the premises is a weakness for libraries, which COVID-19 has made very apparent. The required skill sets among librarians, too, are often lacking to put new technologies into use, thereby making the digital transformation difficult to realize.

1 UP School of Economics

2 UP School of Economics

3 UP School of Library and Information Studies

4 UP Virata School of Business

Evidence and Analysis

Analyzing the 2025 National Library Directory data with Python and Machine Learning classification models gives us important real-world information:

Model	Accuracy	Precision (Weighted)	Recall (Weighted)	F1-Score (Weighted)	Notes
Random Forest	0.727	0.618	0.727	0.668	Performs well overall; strong on predicting Active; class imbalance affects minority classes.
Logistic Regression	0.667	0.726	0.667	0.691	Higher precision; moderately balanced performance.
SVM (RBF Kernel)	0.091	0.731	0.091	0.065	Fails due to severe class imbalance; predicts nearly all cases as Active.
Gradient Boosting	0.788 (Best)	0.708	0.788 (Best)	0.736 (Best)	Strongest model; learns well from sparse features; best generalization.

The results are in line with the main goal of the policy brief because they directly address the most important issue, which is how to use machine learning and AI technologies to improve public libraries and barangay reading centers. The outcomes have made it very clear that there are certain areas, e.g., lack of sufficient ICT infrastructure, low digital literacy, no adaptive tech, and uneven implementation, where RA No. 7743 is lagging behind, and this gives a strong reason for introducing AI and machine learning to the picture. Apart from that, the outcomes also underscore AI's potential and personified tools such as chatbots, adaptive e-learning models, and predictive analytics to enhance personalized learning, improve resource organization, and support community education, which are the libraries' key attributes for becoming data-driven and future-ready learning hubs. Besides, the challenges—privacy issues, lack of infrastructure, high prices, and limited AI literacy - are discussed and provide a very realistic ground for new policy, implying precisely what has to be given attention in order to incorporate machine learning into RA No. 7743. Moreover, the results place more emphasis on the importance of the alignment of community learning centers' modernization with national digitization and Sustainable Development Goal 4 by pointing out that the key issues are the same as those of the poorest, i.e., inclusivity, educational equity, and quality learning outcomes. The findings, all in all, not only back the goal but also push it further, thereby rendering the results in line with the analysis intended direction.

Policy Gaps

A review of RA No. 7743 in light of current data uncovers four significant deficiencies that AI/ML solutions can rectify:

Challenge Area	Current Policy Gap	Proposed AI/ML Solution
Data Monitoring	Reliance on manual data collection and outdated records.	National Library Data Dashboard: Real-time updates and predictive maintenance.
Funding	Uniform allocation regardless of actual library status or need.	Predictive Funding Models: Use ML to direct funds to libraries identified as "at-risk."
Infrastructure	Weak ICT integration, particularly in rural areas.	AI-Enabled Hybrid Hubs: Kiosks providing digital access and learning tools.
Literacy	Limited training for librarians in digital tools.	AI Literacy Programs: Automated training for digital content curation.

Policy Recommendations

In order to update RA No. 7743, the following are recommended:

1. Design the NAELMS. Use ML algorithms in the development of National AI-Enabled Library Monitoring System (NAELMS) to classify libraries in real time as active, at-risk, or closed. This allows for early intervention rather than reactive measures.

2. Institutionalize the "Library Digital Readiness Index." This measure aims to develop an AI-enabled national index that measures infrastructure, connectivity, and learning engagement in all public libraries to meet agreed standards.

3. Ensure Data-Driven Governance Ordinance is law: Include a provision that AI shall assist in annual checking. This means that the amendment of the Implementing Rules and Regulations (IRR) of RA No. 7743 shifts from building libraries to making sure they function.

4. Establish AI Resource Centers in each region. House these alongside thriving provincial libraries to enable their use also as training centers. These will support smaller city centers and encourage a network of innovation rather than stand-alone failures.

5. Convert nonperforming centers into hybrid learning community centers. This aims to provide e-learning services and help people acquire digital skills and access government services at these centers. This will help attain SDG 4 on good education.

Implications and Conclusion

Machine learning in RA No. 7743 changes public libraries from places where people can store information to smart places where people can learn. This will help the government move from a reactive planning strategy to a proactive one. By using their predictions, the government can close the digital divide in a fair and equal way by giving more resources to areas that need them most, like rural and municipal libraries.

For example, the current version of RA No. 7743 is like a map that was printed decades ago; it shows where the roads should be, not where the traffic jams or roadblocks are. Following these AI and ML suggestions turns such a map into a living GPS that will give you real-time information, warn you about problems or closures before they happen, and quickly reroute resources to make sure that every community gets to its goal of digital literacy and lifelong learning.

References

An Act Providing for the Establishment of Congressional, City and Municipal Libraries and Barangay Reading Centers Throughout the Philippines, Appropriating the Necessary Funds Therefor and for Other Purposes. 1994. RA No. 7743 (17 June 1994). https://elibrary.judiciary.gov.ph/thebookshelf/showdocs/2/2399.

National Library of the Philippines. 2025. *Directory of Affiliated Public Libraries in the Philippines*. https://web.nlp.gov.ph/wp-content/uploads/2025/11/OFFICIAL-Integrated-Directory-of-Public-Libraries-as-of-October-2025-FOR-Website-Only.pdf.

Appendix

Dataset Used

The models in the file are trained on the numerical version of the public library dataset (the one with Status as the target and encoded features).

The Status values (Active, Inactive, Closed) encoded as:

```
Active = 1
Inactive = 0
Closed = -1
```

Train/Test Split

The split used in the code is train_test_split(..., stratify=y, test_size=0.2) which matches standard practice.

If the original evaluation also used an 80/20 split and no resampling, the metrics will match.

Model Configurations

The code uses default hyperparameters, which seems to match our report (no tuning mentioned).

If hyperparameter tuning (like GridSearchCV) was used in our paper, the numbers may differ slightly.

How to Confirm?

To verify 100%:

1. Run the code on Complete_Numerical_Conversion_with_Legend.xlsx.
2. Print y.value_counts() to confirm class imbalance (most likely majority = Active).
3. Compare metrics from sklearn.metrics with those in your document.

Four machine learning models were evaluated

1. Random Forest
2. Logistic Regression
3. Support Vector Machine (SVM) with RBF Kernel
4. Gradient Boosting (Best Model)

These were assessed using standard metrics (accuracy, precision, recall, and F1-score) to classify library statuses (e.g., Active, Inactive, Closed) using features from our dataset.

Here is the complete Python code (using scikit-learn) for training and evaluating these models

- Setup and Preprocessing

```
import pandas as pd

from sklearn.model_selection import train_test_split

from sklearn.preprocessing import LabelEncoder

from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score

from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassifier

from sklearn.linear_model import LogisticRegression

from sklearn.svm import SVC
```

- Load your preprocessed dataset (numerical)

```
df = pd.read_excel("Complete_Numerical_Conversion_with_Legend.xlsx", sheet_name="Numerical Data")

# Target variable

y = df["Status_Num"]

# Drop non-feature or redundant columns

drop_cols = ["Status", "Date Affiliated", "Status_Num"]

X = df.drop(columns=[col for col in drop_cols if col in df.columns])

# Train/test split

X_train, X_test, y_train, y_test = train_test_split(X, y, stratify=y, test_size=0.2, random_state=42)
```

- Random Forest Classifier

```
rf = RandomForestClassifier(random_state=42)

rf.fit(X_train, y_train)

y_pred_rf = rf.predict(X_test)


print("Random Forest:")

print("Accuracy:", accuracy_score(y_test, y_pred_rf))

print("Precision:", precision_score(y_test, y_pred_rf, average='weighted'))

print("Recall:", recall_score(y_test, y_pred_rf, average='weighted'))

print("F1-Score:", f1_score(y_test, y_pred_rf, average='weighted'))
```

- Logistic Regression

```
lr = LogisticRegression(max_iter=1000, random_state=42)

lr.fit(X_train, y_train)

y_pred_lr = lr.predict(X_test)


print("Logistic Regression:")

print("Accuracy:", accuracy_score(y_test, y_pred_lr))
```

```
print("Precision:", precision_score(y_test, y_pred_lr, average='weighted'))

print("Recall:", recall_score(y_test, y_pred_lr, average='weighted'))

print("F1-Score:", f1_score(y_test, y_pred_lr, average='weighted'))
```

- SVM with RBF Kernel

```
svm = SVC(kernel='rbf', random_state=42)

svm.fit(X_train, y_train)

y_pred_svm = svm.predict(X_test)


print("SVM (RBF Kernel):")

print("Accuracy:", accuracy_score(y_test, y_pred_svm))

print("Precision:", precision_score(y_test, y_pred_svm, average='weighted'))

print("Recall:", recall_score(y_test, y_pred_svm, average='weighted'))

print("F1-Score:", f1_score(y_test, y_pred_svm, average='weighted'))
```

- Gradient Boosting Classifier (Best Model)

```
gb = GradientBoostingClassifier(random_state=42)

gb.fit(X_train, y_train)

y_pred_gb = gb.predict(X_test)


print("Gradient Boosting:")

print("Accuracy:", accuracy_score(y_test, y_pred_gb))

print("Precision:", precision_score(y_test, y_pred_gb, average='weighted'))

print("Recall:", recall_score(y_test, y_pred_gb, average='weighted'))

print("F1-Score:", f1_score(y_test, y_pred_gb, average='weighted'))
```

Optional: Class Distribution Check (helps explain SVM issue)

```
print("Class Distribution:\n", y.value_counts())
```

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	0.727	0.618	0.727	0.668
Logistic Regression	0.667	0.726	0.667	0.691
SVM (RBF)	0.091	0.731	0.091	0.065
Gradient Boosting	0.788	0.708	0.788	0.736

THE UP CIDS POLICY BRIEF SERIES

The UP CIDS Policy Brief Series features short reports, analyses, and commentaries on issues of national significance and aims to provide research-based inputs for public policy.

Policy briefs contain findings on issues that are aligned with the core agenda of the research programs under the University of the Philippines Center for Integrative and Development Studies (UP CIDS).

Center For Integrative And Development Studies

Established in 1985 by University of the Philippines (UP) President Edgardo J. Angara, the UP Center for Integrative and Development Studies (UP CIDS) is the policy research unit of the University that connects disciplines and scholars across the several units of the UP System. It is mandated to encourage collaborative and rigorous research addressing issues of national significance by supporting scholars and securing funding, enabling them to produce outputs and recommendations for public policy.

The UP CIDS currently has twelve research programs that are clustered under the areas of education and capacity building, development, and social, political, and cultural studies. It publishes policy briefs, monographs, webinar/conference/forum proceedings, and the Philippine Journal for Public Policy, all of which can be downloaded free from the UP CIDS website.

The Program

The Program on Data Science for Public Policy (DSPPP) aims to build the data science knowledge and capacities of academics, researchers, and policymakers, as well as, decision-makers from various sectoral stakeholders, and apply this learned skill to public policy and governance. DSPPP strives to engage a community of researchers within the university and encourage the pursuit of interdisciplinary problem-oriented research using high-level quantitative analysis. It seeks to convene multidisciplinary teams of social scientists, humanists, and scientists to research issues in the public sector.

Get your policy papers published. Download open-access articles.

The Philippine Journal of Public Policy: Interdisciplinary Development Perspectives (PJPP), the annual peer-reviewed journal of the UP Center for Integrative and Development Studies (UP CIDS), welcomes submissions in the form of full-length policy-oriented manuscripts, book reviews, essays, and commentaries. The PJPP provides a multidisciplinary forum for examining contemporary social, cultural, economic, and political issues in the Philippines and elsewhere. Submissions are welcome year-round.

For more information, visit cids.up.edu.ph. All issues/articles of the PJPP can be downloaded for free.

Get news and the latest publications.

Join our mailing list to get our publications delivered straight to your inbox! Also, you'll receive news of upcoming webinars and other updates.

bit.ly/signup_cids

We need your feedback.

Have our publications been useful? Tell us what you think.

bit.ly/dearcids

UNIVERSITY OF THE PHILIPPINES
CENTER FOR INTEGRATIVE AND DEVELOPMENT STUDIES

Lower Ground Floor, Ang Bahay ng Alumni, Magsaysay Avenue
University of the Philippines Diliman, Quezon City 1101

cids.up.edu.ph

Telephone (02) 8981-8500 loc. 4266 to 4268
(02) 8426-0955

Email cids@up.edu.ph
cidspublications@up.edu.ph

Website cids.up.edu.ph