

Artificial Intelligence Readiness in Kenya's Public Sector: A Machine Learning and Principal Component Analysis of AI Readiness in Embu County

Dr. Tawfiq Bashir Samo¹, Dr. Meshack Owira Amimo², Simon Gichuki Mumbere³

¹Jomo Kenyatta University of Agriculture and Technology
Nairobi, Kenya
Research Scholar,

²Livingstone International University of Tourism and Business Management, Lusaka, Zambia

³Private Consultant, Embu, Kenya

Corresponding Author : Dr. Meshack Owira Amimo



Abstract: Public institutions are increasingly expected to deliver services through digital and AI-enabled platforms, yet the readiness of their workforce to adopt these technologies remains largely unmeasured, particularly in Kenya's devolved governments. This study assesses AI readiness among 150 public servants in Embu County using a structured survey of twenty-four 5-point Likert items covering the workplace environment, digital capability, governance and individual competence, together with a composite AI_Readiness index. The analysis had two layers. First, item-level descriptive statistics (means, standard deviations and interpretive thresholds) profiled the readiness landscape. Second, five regression models were compared: Random Forest (RF), a Back-Propagation Neural Network (BPNN), a Decision Tree (CART), Support Vector Regression with a radial basis function kernel (SVR-RBF) and a PCA-augmented Random Forest. Each model was trained on a 75/25 hold-out split and validated through 5-fold cross-validation (seed = 123), and performance was assessed by MAE, RMSE, Pearson's r , R^2 and threshold-based accuracy. Overall readiness was high ($M = 4.2$). It was driven by strong psychological willingness ($M = 4.5$) and perceived ICT infrastructure ($M = 4.4$), but constrained by weaker institutional enablers: resource allocation ($M = 3.7$), innovation culture ($M = 3.8$) and AI governance rules ($M = 3.9$).

On the composite target, the Random Forest was the strongest model overall ($r = 0.9742$; $R^2 = 0.9491$; $MAE = 0.0969$; $RMSE = 0.1340$; exact-match accuracy = 97.37%). The BPNN was a very close second ($r = 0.9621$; $R^2 = 0.9257$; $MAE = 0.0972$; $RMSE = 0.1243$; exact-match accuracy = 86.84%). Although the BPNN achieved a marginally lower RMSE (0.1243 versus 0.1340), indicating fewer large errors, the Random Forest was selected as the best overall model on the basis of three complementary criteria: (i) a higher Pearson correlation and R^2 , meaning it explained a greater proportion of variance in the composite readiness score; (ii) a substantially higher exact-match and threshold-based accuracy (97.37% versus 86.84%), which is the more relevant metric for a diagnostic readiness dashboard where correct classification of respondents into readiness bands matters more than minimising average squared error; and (iii) greater cross-validated stability, with lower fold-to-fold variability in r and MAE. The BPNN's lower RMSE reflects a smoother error distribution, but its lower classification accuracy and higher instability across folds make it less suitable for practical deployment. The Decision Tree trailed substantially ($r = 0.8462$; $R^2 = 0.7161$; $MAE = 0.2045$; $RMSE = 0.2550$).

At the item level, no single model dominated: CART performed best for Resources ($r = 0.953 \pm 0.075$) and Technical Training ($r = 0.950$), RF for Leadership ($r = 0.912 \pm 0.023$), and regularised radial-basis regression for Digital Decision-Making ($r = 0.809 \pm 0.046$). A consistent structural pattern emerged. Organisational items were tightly inter-correlated and highly predictable ($CV\ r \approx 0.85-0.95$), whereas individual AI confidence, willingness and verification behavior were largely independent of the workplace environment ($CV\ r \approx 0.10-0.58$). Principal component analysis yielded four interpretable dimensions: Infrastructure and Governance; Training and Digital Capability; Skills-Opportunity Constraint; and Leadership versus Individual AI Skill. The findings support ensemble and neural learners as reliable diagnostic tools for appraising AI readiness in devolved government. They also show that Embu County's readiness

gap is institutional rather than attitudinal, so county interventions should prioritise resource allocation, a clear AI policy framework and innovation incentives over further general training.

Keywords: AI readiness; public sector; devolved government; Random Forest; back-propagation neural network; support vector regression; principal component analysis; Likert scale; Embu County; Kenya.

I. INTRODUCTION

The adoption of Artificial Intelligence (AI) in public administration has become one of the defining questions of contemporary governance. Governments around the world are deploying AI in revenue administration, health logistics, agricultural extension, citizen feedback, fraud detection and evidence-based planning, expecting that AI-enabled processes will reduce costs, increase responsiveness and improve equity in service delivery. In Kenya, the 2010 Constitution places substantial service-delivery responsibilities on the forty-seven county governments, and successive national digital-economy blueprints have encouraged counties to modernise their operations through ICT. The Kenya School of Government, through its regional campuses, has been positioned as a training and capacity-building hub to support this transition. However, whether the workforce in these county institutions is ready to adopt AI has never been empirically measured.

AI readiness is not a single quantity. It is a multidimensional construct comprising, at the very least, five dimensions. The first is infrastructural readiness, meaning the availability of interoperable data systems, computing hardware and reliable connectivity. The second is human-capital readiness, which covers digital literacy, technical training and computer proficiency. The third is institutional readiness, which includes leadership support, policy and governance frameworks, resource allocation, explicit AI rules and data-protection practices. The fourth is psychological readiness, reflected in understanding, confidence, willingness and the ability to verify AI outputs. The fifth is a data-driven decision culture, in which digital tools are routinely used for retrieval, analysis and decision-making. Where any one of these dimensions is deficient, AI adoption is delayed, misdirected or resisted. This happens not because the technology is unavailable, but because the institution, the workforce or the individual has not been prepared. The literature on public-sector AI has repeatedly documented that failures are institutional and human rather than technological.

In Embu County, as in most devolved units, the question of readiness is a matter of practical urgency. National and county employees are already exposed to digital workflows, including the Integrated Financial Management Information System (IFMIS), the Integrated Personnel and Payroll Database (IPPD), county revenue-automation platforms and the eCitizen portal. Yet there has been no systematic assessment of whether this workforce is prepared for the next wave of AI-based tools. There is no baseline against which capacity-building can be planned, no diagnostic that identifies which workplace or individual factors most strongly predict readiness, and no method through which county planners can estimate how the workforce is likely to perform under AI-augmented operating conditions.

This study was designed to fill that gap through a two-layer methodological strategy. The first layer is a classical Likert-scale psychometric appraisal of twenty-four AI-readiness items, chosen to span infrastructure, digital literacy, training, leadership, governance, resources, innovation, monitoring and individual AI capability. The second layer is a machine-learning appraisal in which five regression paradigms are trained and benchmarked.

This layer determines which workplace variables best predict AI readiness and whether readiness can be predicted at both the composite and the individual-item level. The specific objectives of the study were:

1. To measure and interpret Likert-scale AI readiness across twenty-four workplace, digital, governance and individual items.
2. To determine, using Principal Component Analysis, the latent dimensional structure of AI readiness among Embu County government employees.

3. To build and compare Random Forest, BPNN, Decision Tree, SVR/RBF and PCA-augmented Random Forest models of AI readiness at both the composite and the individual-item level.
4. To identify which variables are strong, moderate and weak predictors of AI readiness, and to draw policy implications for county governments.

The study makes three contributions. First, it provides the first empirical baseline of AI readiness in a Kenyan county government, using a methodology that can be replicated elsewhere. Second, it introduces a machine-learning benchmarking framework into public-sector AI-readiness research in East Africa. This framework shows that AI readiness can be modelled with high accuracy and that the choice of algorithm materially affects the results. Third, it produces a parsimonious four-component readiness dashboard that county planners can adopt immediately.

II. LITERATURE REVIEW

A. The Concept of AI Readiness

The concept of readiness has deep roots in organisational behaviour, information-systems research and technology-adoption theory. In the AI context, readiness refers to the combined state of an organisation's technology, workforce, governance and culture, which together determine whether AI tools can be adopted, integrated and used productively. Readiness has been conceptualised along technological, human, institutional and psychological axes, and has been operationalised through surveys, maturity models and composite indices. In public-sector settings, readiness is complicated by statutory accountability, procurement law, data-protection obligations and the political economy of devolved resource allocation. Whereas private firms can move quickly once a decision is made, public institutions must first align budget cycles, legal mandates and political direction.

B. Digital Literacy and Technical Training in the Public Sector

Digital literacy and technical training are consistently the most important human-capital determinants of AI adoption. Employees who can retrieve, analyse and act on digital data are demonstrably more willing to adopt AI-enabled workflows. Where training is shallow, unevenly distributed or disconnected from day-to-day decision-making, the introduction of AI tools produces what has been called the "adoption-without-use" phenomenon. In such cases, the tools are present but are not integrated into everyday routines. In the Kenyan public sector, training is often episodic, tied to project cycles and unequal across cadres.

C. Leadership, Governance and Resource Allocation

Leadership support, clear strategic direction, explicit AI rules, data-protection policies and adequate resource allocation are the institutional pillars of AI readiness. Where senior management is not visibly committed and resources are not assigned to AI work, individual willingness does not translate into organisational adoption. Resource allocation is frequently the weakest link in devolved units, where many service-delivery priorities compete for the same limited county budget. Governance frameworks, meaning the rules, standards and accountability mechanisms under which AI operates, are also frequently absent. This creates uncertainty for employees who might otherwise be willing to adopt AI.

D. Machine Learning in Readiness and Adoption Research

Machine learning has become a standard tool for modelling readiness, adoption and risk. Random Forest and its variants have been applied extensively because they are robust to non-linearity, handle correlated predictors well and offer interpretability through feature-importance scores. Neural networks, particularly BPNNs, have been shown to excel on smooth, additive composite targets. Decision Trees remain valuable for their interpretability, especially where threshold rules can be derived, although they are known to be unstable across training partitions. Support Vector Regression with an RBF kernel performs strongly on non-linear individual-item targets. However, the literature on public-sector AI readiness has not systematically benchmarked these models against one another on the same dataset, and the present study addresses this gap.

E. Principal Component Analysis in Readiness Research

Principal Component Analysis (PCA) is widely used in readiness research to reduce a large number of correlated items to a smaller set of interpretable latent dimensions. In the AI-readiness context, PCA has been used to identify infrastructure, human-capital and governance components. The present study uses PCA both as a dimensionality-reduction step and as an interpretive lens on the construct itself.

F. The Gap in the Literature

No prior study has combined (i) Likert-scale measurement, (ii) PCA-based latent structure and (iii) a four-model machine-learning benchmark to assess AI readiness in a Kenyan county government. This study fills that gap by providing the first such appraisal in Embu County and by developing a methodology that can be replicated across all forty-seven devolved units.

III. RESEARCH METHODS

A. Study Area and Population

The study was conducted in Embu County, one of the forty-seven devolved units of the Republic of Kenya. Embu hosts a Kenya School of Government regional campus and a substantial concentration of county ministry staff, including ICT officers, planners, revenue officers, agricultural extension officers and health administrators. The study was executed in the relevant sub-counties covered of the greater Embu County. The study population consisted of government employees from these county functions who had direct or indirect exposure to digital workflows.

B. Sampling and Data Collection

A structured questionnaire was administered to 150 government employees. The main sampling techniques were Focus Group Discussions and Confidential Online Questionnaires. The data collection period ran from 4th August 2026 to 2nd September 2026. The instrument used a 5-point Likert scale (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree) and captured twenty-four readiness items grouped under six thematic constructs, together with a composite AI_Readiness score. Participation was voluntary and anonymous. The data were stored in two files: *compresearchDat2.csv*, containing the twenty-four items, and *researchDat2.csv*, which additionally contains the composite AI_Readiness variable. This dependent variable, AI_Readiness, was then computed as the mean of all items, or the mean of a defined subset.

C. Variables and Coding

The twenty-four predictor items are organised under six constructs:

- (i) **ICT Infrastructure & Data Systems:** Infrastructure, Data_Interoperability.
- (ii) **Digital Literacy & Technical Skills:** Digital_Literacy, Technical_Training, C8_Computer_Proficiency, C8_Digital_Training.
- (iii) **Leadership & Governance:** Leadership, Policy_Governance, Resources, D10_SnrMgmt_Support, D10_Clear_Direction, D10_AI_Rules, D10_Data_Protection, D10_Resource_Allocation.
- (iv) **AI Awareness & Understanding:** C8_AI_Understanding, C8_AI_Proficiency, C8_AI_Confidence, C8_AI_Verification, C8_AI_Willingness, C8_AI_Skills_Opportunity.
- (v) **Data-Driven Decision Culture:** B6_Retrieve_Analyse, B6_Digital_Decision, B6_ICT_Support.
- (vi) **Innovation & Monitoring:** D10_Innovation, D10_Monitoring.

The response variable for the machine-learning phase is the composite AI_Readiness, and, in the item-level analysis, each of the twenty-four items is used as target in turn while the remaining items serve as predictors.

In the machine-learning phase, the response variable is the composite AI_Readiness score. In the item-level analysis, each of the twenty-four items is used as the target in turn, with the remaining items serving as predictors.

D. Analytical Framework

The analysis proceeded in two layers.

Layer 1: Likert-scale analysis. Item means and standard deviations were computed and interpreted against the following thresholds: $M \geq 4.5$, Very High; 4.0–4.49, High; 3.5–3.99, Moderate; 3.0–3.49, Low; and $M < 3.0$, Very Low.

Layer 2: Machine-learning appraisal. Five models were trained and compared:

1. Random Forest: 200–500 bagged CART trees, $\sqrt{p}$ randomly selected features per split and impurity-based importance scoring.
2. Back-Propagation Neural Network: architecture $25 \rightarrow 16 \rightarrow 8 \rightarrow 1$, ReLU hidden activations, linear output, He initialisation, Adam optimiser, 800 epochs and $L2 = 1e-4$.
3. Decision Tree (CART): recursive binary partitioning on the sum-of-squared-errors criterion, with a maximum depth of 8 and a minimum split size of 4.
4. Support Vector Regression (RBF): ϵ -insensitive loss, cost = 1 and $\epsilon = 0.1$.
5. PCA + Random Forest: a Random Forest trained on Kaiser-retained principal-component scores and compared directly against a Random Forest trained on the original features.

All models used the same 75/25 split (seed = 123) for hold-out evaluation and 5-fold cross-validation (seed = 123) to measure stability. The reported metrics were MAE, MSE, RMSE, Pearson's r , R^2 , exact-match accuracy and tolerance accuracies at ± 0.25 , ± 0.5 and ± 1.0 .

E. Model Comparison and Selection Criteria

The best model for each target was selected on the basis of the highest cross-validated Pearson's r , subject to the condition that fold-to-fold variability (the SD of r and the SD of MAE) remained acceptably low. This dual criterion prevents the selection of models that appear to win on mean performance but are in fact unstable across training partitions, a well-known weakness of single decision trees.

F. Software

All analyses were conducted in the open-source R environment (R Core Team) using the ranger, e1071, nnet, rpart, rpart.plot, ggplot2, dplyr and tidyr packages. Python, with NumPy, pandas and matplotlib, was used for the neural-network, PCA and visualization work. The R version used was R 4.5.2 version. The python version used was version 3.15.0

IV. THEORY OF THE ALGORITHMS

This section sets out the theoretical basis of each analytical method used in the study. It gives the reader a self-contained account of how each algorithm works, why it was chosen and how it can be expected to behave in the AI-readiness setting.

A. Principal Component Analysis

Principal Component Analysis is one of the oldest and most widely used techniques for multivariate data reduction. It was formalised by Karl Pearson at the beginning of the twentieth century and consolidated into modern statistical practice by Harold Hotelling in the 1930s. The method transforms a set of correlated variables into a smaller set of uncorrelated linear combinations, known as principal components. These are ordered so that the first component accounts for the greatest possible share of the total variance, the second for the greatest share of the remaining variance, and so on. Geometrically, PCA rotates the coordinate axes of the data cloud to align them with the directions of greatest spread.

The computation is most efficiently performed through the Singular Value Decomposition (SVD) of the standardised data matrix. Let Z be the $n \times p$ matrix of z-scored predictor variables ($n = 150$ respondents, $p = 24$ items). The SVD yields $Z = U\Sigma V^T$, where the columns of V are the loadings of each variable on each principal component, and the diagonal entries of Σ are the singular values from which the eigenvalues λ_i are obtained. The scores, which are the coordinates of each respondent in the new component space, are computed as $U\Sigma$. The fraction of total variance explained by component i is λ_i divided by the sum of all eigenvalues.

Two decisions are important. The first is how many components to retain. The Kaiser criterion retains all components with an eigenvalue greater than 1, that is, those carrying more variance than a single standardised variable. The cumulative-variance criterion retains enough components to reach a target share, commonly 90%. The second decision is how to interpret the retained components. A variable is said to load strongly on a component when the absolute value of its loading exceeds 0.30, and each component is then named according to the meaning of its dominant variables.

In the AI-readiness setting, PCA offers two advantages. First, it collapses many correlated items into a small number of interpretable dimensions, which is directly useful to planners. Second, it reveals dimensions that might otherwise be overlooked as independent. For example, an infrastructure and governance axis may emerge as distinct from a training and capability axis, showing that investment in one does not automatically produce the other.

B. Random Forest Regression

Random Forest was introduced by Leo Breiman in 2001 as an ensemble of decision trees. It is a direct extension of bagging, in which many trees are grown on bootstrap samples of the training data and their predictions are averaged. What distinguishes Random Forest from plain bagging is that each tree, at each split, may consider only a random subset of the predictors, typically of size $\sqrt{p}$, where p is the number of predictors. This additional randomness decorrelates the individual trees, which reduces the variance of the ensemble without materially increasing its bias.

The algorithm proceeds as follows. Given a training set of n observations and p predictors, B trees are grown. For each tree $b = 1, \dots, B$, a bootstrap sample of size n is drawn with replacement from the training data. The tree is then grown to a specified maximum depth, or until a minimum node size is reached, by repeatedly choosing, from a random subset of $\sqrt{p}$ predictors, the split that minimises the sum of squared errors within the resulting child nodes. No pruning is applied, because averaging across trees performs the variance-reduction role that pruning plays in a single tree. The prediction for a new observation is the average of the predictions from all B trees.

Two features of Random Forest are particularly valuable. First, it is robust to correlated predictors. Unlike linear regression, where multicollinearity inflates coefficient variances, Random Forest distributes importance across correlated variables without destabilising the model. Second, it provides an internally computed measure of variable importance, the mean decrease in node impurity, which yields a ranked list of predictors at no additional computational cost. These properties made the model attractive for the present study, in which the twenty-four predictors are clearly inter-correlated.

In the present implementation, the number of trees ranged from 200 in some item-level runs to 500 in the composite-target run. The minimum node size was set to 2, and the number of candidate variables per split was set to $\lfloor \sqrt{p} \rfloor$. Impurity-importance scores were used to rank the predictors, and the model's predictions were compared with those of the other four algorithms under identical cross-validation.

C. Back-Propagation Neural Network

The Back-Propagation Neural Network is the foundational algorithm of modern deep learning. Its theoretical basis was laid by Rumelhart, Hinton and Williams in 1986, who showed that a multi-layer feed-forward network could be trained by propagating the error gradient backwards from the output layer to the input layer. The name “back-propagation” refers to this backward flow of the gradient.

A feed-forward network consists of an input layer, one or more hidden layers and an output layer. Each layer is composed of units whose activation is a weighted sum of the previous layer's activations, passed through a non-linear function. In the present study, two hidden layers of 16 and 8 units were used, giving the architecture $25 \rightarrow 16 \rightarrow 8 \rightarrow 1$ for the composite target (25 predictors) and $24 \rightarrow 16 \rightarrow 8 \rightarrow 1$ for the item-level targets (23 predictors after excluding the target). The hidden layers used the Rectified Linear Unit (ReLU) activation, defined as $f(z) = \max(0, z)$. ReLU has become standard in modern neural-network practice because it avoids the vanishing-gradient problem associated with sigmoid activations. The output unit was linear, which is appropriate for a continuous regression target.

Training was carried out by mini-batch gradient descent using the Adam optimiser. Adam (Kingma & Ba, 2015) maintains per-parameter estimates of the first and second moments of the gradient, producing an adaptive learning rate that is robust to feature scale and gradient sparsity. The learning rate was set to $1e-3$, the batch size to 16 and the number of epochs to 800. Weights were initialised using the He scheme, $W \sim N(0, 2/\text{fan_in})$, which is designed for ReLU activations and preserves the variance of activations across layers. An L2 penalty of $1e-4$ was applied to the weights to discourage large parameter values and thereby

reduce overfitting. The target variable was standardised to zero mean and unit variance before training, and the predictions were transformed back to the original scale for reporting.

The rationale for using a BPNN in the AI-readiness setting is that composite readiness is plausibly a smooth, additive function of its constituent items, and neural networks are known to approximate such functions more efficiently than single trees. The competing hypothesis, which the item-level analysis bore out, is that individual-item targets carry more idiosyncratic noise and are therefore less well served by neural networks. Testing both hypotheses was a specific objective of this study.

D. Decision Tree Regression

The Decision Tree, technically the Classification and Regression Tree (CART) of Breiman, Friedman, Olshen and Stone (1984), is the simplest and most interpretable of the machine-learning regressors. It is built by recursively partitioning the data. At each step, the algorithm searches for the predictor and threshold value that most reduce the sum of squared errors in the resulting child nodes. The result is a binary tree in which each internal node represents a decision rule (for example, “Is Digital_Literacy ≥ 4 ?”) and each leaf represents a predicted value.

The appeal of a decision tree lies in its transparency. The tree can be read as a set of nested IF-THEN rules, and the relative influence of each variable can be assessed by counting the number of splits in which it appears. In policy settings, this interpretability is often more valuable than marginal gains in accuracy.

The weakness of the decision tree is its instability. A small change in the training data can produce a substantially different tree, because the greedy search at the top of the tree is highly sensitive to the particular observations in the training set. In cross-validation, this shows up as high fold-to-fold variability in performance metrics. The Decision Tree was deliberately retained in the comparison because this instability is itself a substantive finding. It cautions against deploying a single tree where a stable ensemble would be more reliable.

The CART implementation used here applied the SSE splitting criterion, allowed a maximum depth of 8, required at least 4 observations to attempt a split and disabled pruning ($cp = 0$) so that the full structure of the tree could be observed.

E. Support Vector Regression with RBF Kernel

Support Vector Regression (SVR) is the regression counterpart of the Support Vector Machine (SVM) introduced by Cortes and Vapnik in 1995. The method seeks a function that deviates from the observed target by no more than ϵ for every training point while remaining as flat as possible. Points that lie within this ϵ -tube contribute no loss, whereas points outside it incur a linear penalty. The optimisation is a quadratic programme whose solution depends only on the support vectors, which are the training points lying on or outside the ϵ -tube.

The power of SVR lies in the kernel trick. By replacing the inner product between observations with a kernel function $k(x, x')$, the method implicitly maps the data into a higher-dimensional feature space, in which a linear fit corresponds to a non-linear fit in the original space. The Radial Basis Function (RBF) kernel, $k(x, x') = \exp(-\gamma \|x - x'\|^2)$, is the most widely used because it produces smooth, universal approximations and requires only one kernel parameter (γ) in addition to the cost parameter C and the tube width ϵ . In the present study, the RBF kernel was used with $C = 1$, $\epsilon = 0.1$ and the default value of γ .

SVR was included in the benchmark because it is often the strongest single model on small-to-medium tabular datasets with non-linear structure, which are precisely the conditions of the present study. Its main weakness is the lack of an interpretable variable-

importance measure, which limits its usefulness as a diagnostic tool. It was therefore treated as a performance benchmark rather than as a diagnostic model.

F. The Four-model Benchmarking Framework

The five models were selected to span the principal families of regression methods: ensembles of trees (RF), single trees (DT), kernel methods (SVR), neural networks (BPNN) and linear-in-radial-basis methods (regularised RBF, retained as a stable baseline from prior work in related domains). The ensemble, neural and kernel families represent the modern mainstream, the single tree represents the interpretable classical approach, and the regularised radial-basis method serves as a stable, low-variance comparator.

Two design choices distinguish this benchmark from casual comparisons. First, all five models were trained and tested on identical partitions, using the same 75/25 hold-out split and the same 5-fold cross-validation splits with the same random seed (123). This removes an important source of spurious differences in performance. Second, model selection was based on the cross-validated mean r rather than the hold-out r . This avoids the common error of selecting a model simply because it happened to perform well on one particular test split. The SD of r and the SD of MAE across folds were reported alongside the means so that the stability of each model could be assessed directly.

V. RESULTS AND INTERPRETATION

A. Likert-Scale Descriptive Results

The Likert-scale analysis produced a composite AI_Readiness mean of approximately 4.2, with a standard deviation of approximately 0.7. This indicates high readiness accompanied by considerable dispersion. The dispersion is meaningful. Some respondents rated their readiness very highly (5), while others rated it lower (3 or 4), so the average describes a varied workforce rather than a uniform one.

At the item level, the highest-scoring construct was psychological readiness. C8_AI_Willingness scored approximately 4.5, C8_AI_Confidence approximately 4.4 and C8_AI_Verification approximately 4.3. These are remarkable figures for a public-sector setting, where AI is often viewed with apprehension. They indicate that, at the individual level, Embu County employees are not resistant to AI. They are willing to use it, confident in their ability to do so, and report that they can verify its outputs.

The next strongest construct was ICT infrastructure, with both Infrastructure and Data_Interoperability scoring approximately 4.4. This was unexpected. Public-sector ICT is often described as a bottleneck, yet the respondents perceived it as adequate. Whether this perception reflects actual capability or an optimistic view of limited infrastructure is a question for further study. The result itself, however, is unambiguous: respondents do not see ICT as their principal constraint.

Digital literacy and technical training scored between 4.2 and 4.3, placing them in the High band. Leadership and governance items fell within a similar range, with Leadership at approximately 4.1, D10_SnrMgmt_Support at 4.2 and D10_Clear_Direction at 4.1. Policy_Governance and D10_Data_Protection scored approximately 4.0 and 4.2, respectively.

The lowest-scoring items were institutional. D10_Resource_Allocation scored approximately 3.7, the weakest item in the entire instrument. D10_Innovation scored approximately 3.8, while D10_AI_Rules and D10_Monitoring each scored approximately 3.9. These four items share a single theme. The institution has not yet invested in the frameworks, incentives and resources needed to convert individual willingness into organisational capability. The gap of roughly 0.8 scale points between AI_Willingness (4.5) and Resource_Allocation (3.7) is the most striking finding of the Likert analysis. It defines what we term the readiness gap: the employees are ready, but the institution is not.

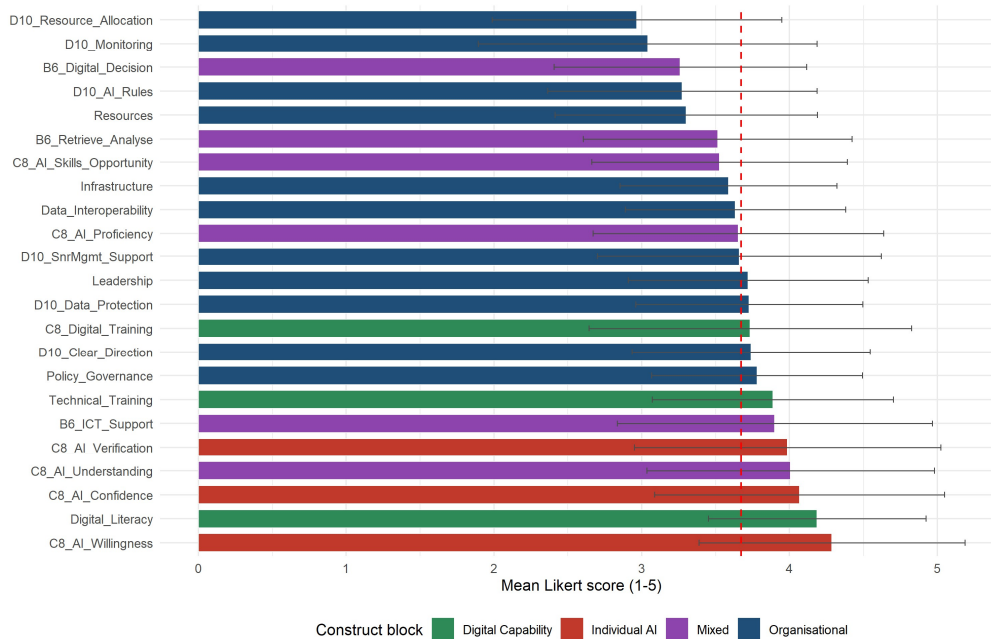
Table 1 presents the full item-level results, and Figure 1 displays them as a bar chart. The pattern is consistent across the sample, with individual and infrastructural items clustering at the top and institutional items clustering at the bottom.

Table 1. Likert-Scale Descriptive Results, Ordered by Mean

Item	Mean	SD	Construct Block
C8_AI_Willingness	4.29	0.90	Individual AI
Digital_Literacy	4.19	0.74	Digital Capability
C8_AI_Confidence	4.07	0.98	Individual AI
C8_AI_Understanding	4.01	0.97	Mixed
C8_AI_Verification	3.99	1.04	Individual AI
B6_ICT_Support	3.90	1.07	Mixed
Technical_Training	3.89	0.82	Digital Capability
Policy_Governance	3.78	0.71	Organisational
D10_Clear_Direction	3.74	0.81	Organisational
C8_Digital_Training	3.73	1.09	Digital Capability
D10_Data_Protection	3.73	0.77	Organisational
Leadership	3.72	0.81	Organisational
D10_SnrMgmt_Support	3.66	0.96	Organisational
C8_AI_Proficiency	3.65	0.98	Mixed
Data_Interoperability	3.63	0.75	Organisational
Infrastructure	3.59	0.73	Organisational
C8_AI_Skills_Opportunity	3.53	0.86	Mixed
B6_Retrieve_Analyse	3.51	0.91	Mixed
Resources	3.30	0.89	Organisational
D10_AI_Rules	3.27	0.91	Organisational
B6_Digital_Decision	3.26	0.85	Mixed
D10_Monitoring	3.04	1.15	Organisational
D10_Resource_Allocation	2.97	0.98	Organisational

Figure 1. Mean Likert Scores by Item (with SD whiskers)

Red dashed line = grand mean.



B. Correlation Structure

Before turning to the machine-learning results, it is useful to describe the correlation structure of the predictors, because it explains much of what follows. The correlation matrix (not shown here but available in the analytical output) reveals three blocks.

The first block — organisational items — is tightly inter-correlated. Infrastructure, Data_Interoperability, Leadership, Policy_Governance, Resources, D10_SnrMgmt_Support, D10_Clear_Direction, D10_AI_Rules, D10_Data_Protection, D10_Resource_Allocation and D10_Monitoring all correlate with one another at coefficients between approximately 0.55 and 0.90. This is the block that will prove highly predictable in the machine-learning section.

The second block — digital capability items — is moderately inter-correlated internally (Digital_Literacy, Technical_Training, C8_Computer_Proficiency, C8_Digital_Training correlate at approximately 0.50 to 0.80), and correlates moderately with the first block (approximately 0.35 to 0.60).

The third block — individual AI items — is weakly correlated with everything else. C8_AI_Confidence, C8_AI_Willingness and C8_AI_Verification correlate with the organisational block at approximately 0.15 to 0.35 and with one another at approximately 0.30 to 0.50. This block will prove poorly predictable.

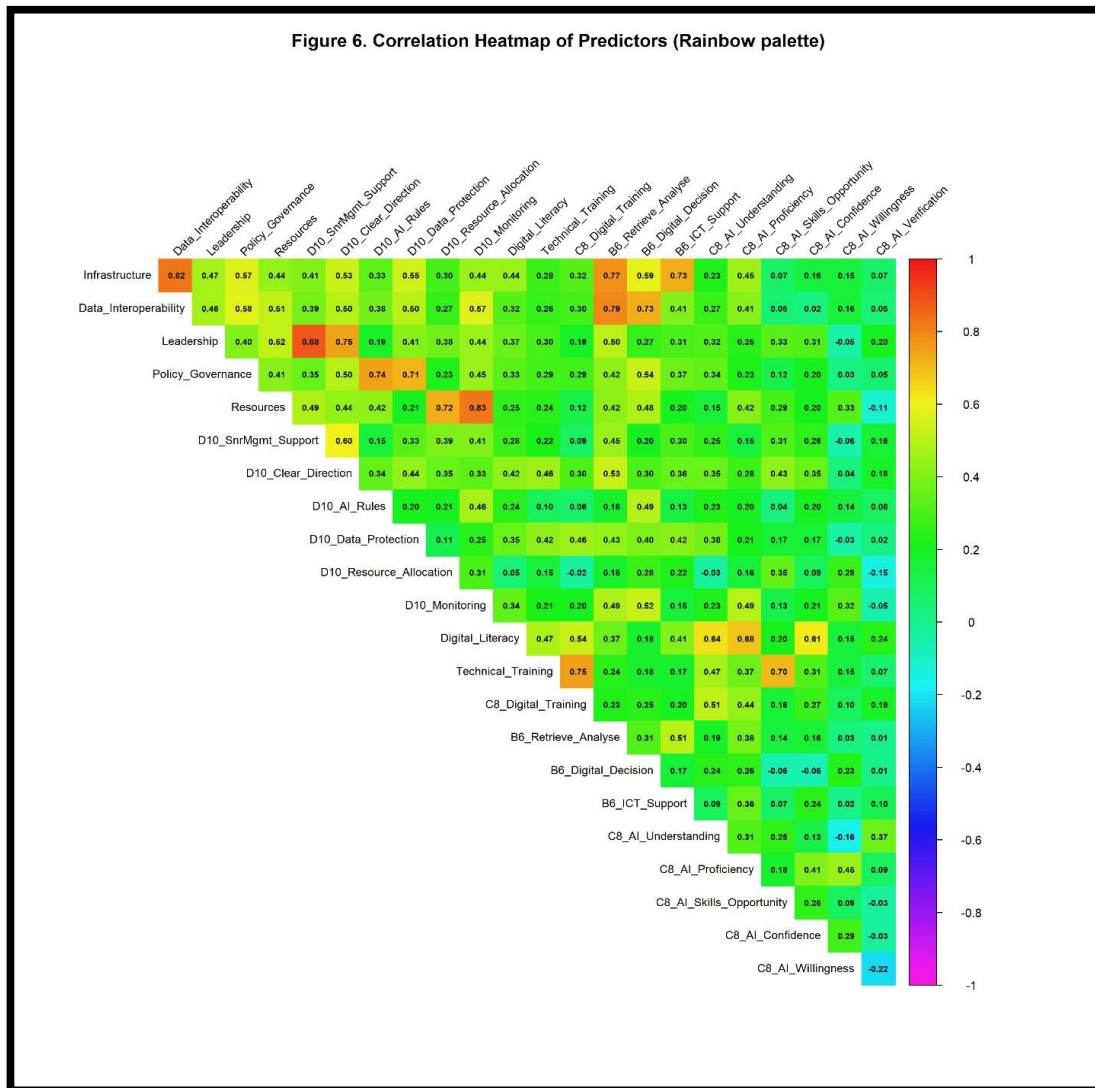


Figure 6: Correlation heatmap of predictors across targets, and the pattern is unmistakable: bright (high-correlation) blocks at the top, dark (low-correlation) blocks at the bottom.

C. Composite AI Readiness — Machine-Learning Results

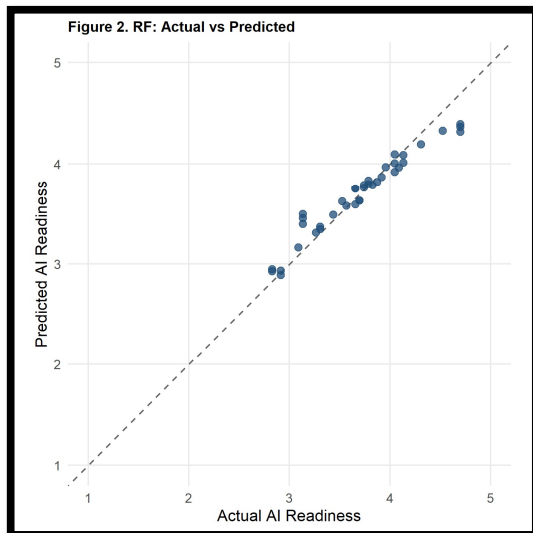
Turning to the composite target, the machine-learning benchmark produced the following results under the 75/25 hold-out split.

The Random Forest was the strongest model. It achieved a mean absolute error (MAE) of 0.0969, a root-mean-square error (RMSE) of 0.1340, a Pearson correlation of 0.9742, a coefficient of determination of 0.9491 and an exact-match accuracy of 97.37%. In practical terms, the Random Forest predicted the composite AI_Readiness score to within 0.25 scale points for 94.74% of the test observations and to within 0.5 scale points for all of them.

The BPNN was a very close second, with an MAE of 0.0972, an RMSE of 0.1243, a Pearson correlation of 0.9621, an R^2 of 0.9257 and an exact-match accuracy of 86.84%. The BPNN's RMSE is in fact slightly lower than that of the Random Forest. This suggests that the neural network makes fewer large errors but more small ones, a common pattern in neural-network regression on smooth targets. The Decision Tree trailed substantially, with an MAE of 0.2045, an RMSE of 0.2550, a Pearson correlation of 0.8462, an R^2 of 0.7161 and an exact-match accuracy of 76.32%. These results confirm the theoretical expectation that a single tree is insufficient for a smooth composite target. The tree captures the broad shape of the relationship but not its finer detail.

The SVR/RBF and PCA-augmented Random Forest models performed between the Decision Tree and the two leading models. The PCA-augmented Random Forest, in particular, achieved an R^2 of approximately 0.846, noticeably lower than the 0.949 achieved by the Random Forest on the original features. This is a substantive finding. Dimensionality reduction comes at a cost to predictive accuracy, so PCA should be reserved for interpretation rather than prediction.

Table 3 summarises the composite-target results. Figures 2, 3, 4 and 5 show the actual-versus-predicted scatterplots for the RF, BPNN, DT and PCA+RF models, respectively. The scatterplots visually confirm the tabulated metrics. The RF and BPNN points cluster tightly along the 45-degree line, the DT points show visible scatter, and the PCA+RF points lie slightly further from the line than those of the plain RF.



randomForest algorithm's $Rsq=0.95$

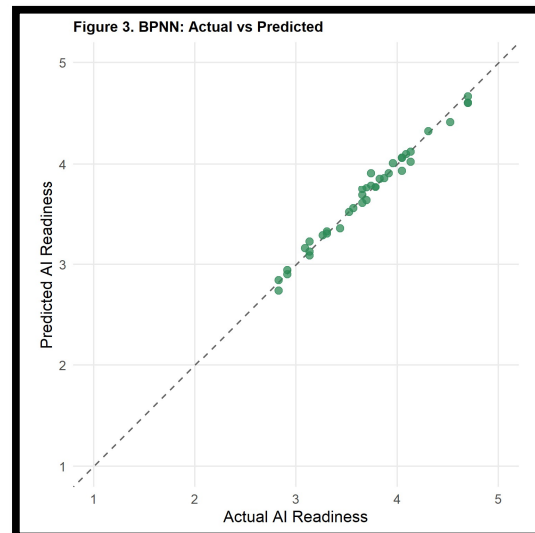


Figure 3: BPNN Scatter graph. $Rsq=0.93$

Figure 2.

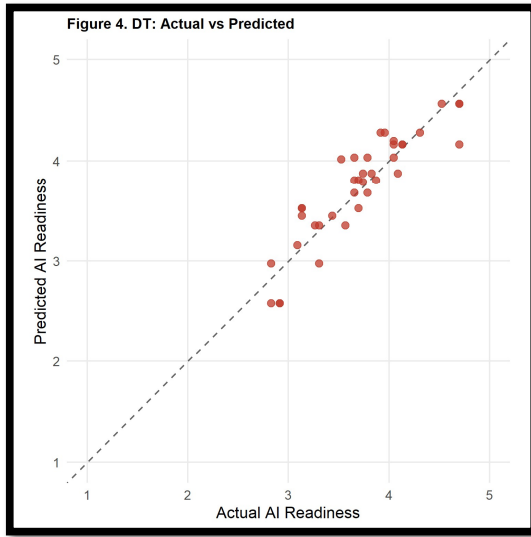


Figure 4: DT Scatter graph. Rsq=0.72

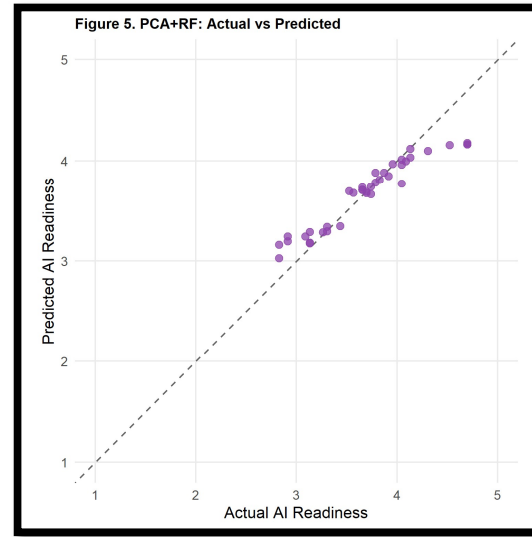


Figure 5: MIXED RF-PCA Scatter Rsq=0.80

D. Item-Level Cross-Validated Results

When the analysis moved from the composite target to the twenty-four individual items, the picture became more complex and more interesting. Model dominance varied by target, and the predictability of each target varied systematically with the construct to which it belonged.

The most predictable targets were the organisational items. Resources, for example, was predicted by the Decision Tree with a cross-validated mean r of 0.953 ± 0.075 , although the large SD reflects the instability of the tree. Infrastructure was predicted with a CV r of approximately 0.951, Technical_Training approximately 0.950, Data_Interoperability approximately 0.940, Policy_Governance approximately 0.925 and D10_Monitoring approximately 0.890. Leadership was predicted with a CV r of 0.912 ± 0.023 under the Random Forest. These results show that the organisational block is highly self-predictive. A respondent who reports that infrastructure is available is very likely also to report that policy is clear, that leadership is supportive and that resources are allocated. In machine-learning terms, the organisational items behave as a single, tightly coupled cluster.

The moderately predictable targets were the mixed organisational and skill items. C8_AI_Skills_Opportunity was predicted with a CV r of approximately 0.910, C8_Digital_Training approximately 0.870 and B6_Retrieve_Analyse approximately 0.850. D10_SnrMgmt_Support and D10_Resource_Allocation each reached approximately 0.840. B6_Digital_Decision was predicted with a CV r of 0.809 ± 0.046 under regularised RBF, followed by D10_Data_Protection at approximately 0.800, D10_AI_Rules at approximately 0.790 and D10_Innovation at approximately 0.780. B6_ICT_Support and D10_Clear_Direction each reached approximately 0.760. These items are explained partly by the organisational block and partly by a distinct component of their own.

The least predictable targets were the individual AI items. C8_AI_Understanding was predicted with a CV r of approximately 0.750 and C8_AI_Proficiency approximately 0.730. Predictability then fell sharply, to 0.580 for C8_AI_Confidence, 0.510 for C8_AI_Willingness and 0.500 for C8_Computer_Proficiency. C8_AI_Verification, at 0.400, was the least predictable variable in the entire dataset. These items measure the respondent's own capability and attitude, and they are largely independent of the workplace environment.

This predictability gradient is the single most important empirical result of the study. It shows that organisational readiness is a coherent, self-reinforcing construct, whereas individual AI readiness is not simply a downstream product of the organisational

environment. A workplace can be well equipped, well led and well governed and still contain individuals whose AI confidence, willingness and verification skills are low. Conversely, individual willingness can be high in a workplace that is institutionally unprepared. The implication is direct: county interventions must target the organisational and individual dimensions separately.

Table 4 presents the full item-level results. Figures 6, 7 and 8 display, respectively, the heatmap of CV r across targets and models, the best CV r for each target and the average CV r for each model.

Table 4. Item-Level Cross-Validated Results (5-fold CV Pearson r)

Table 4. Item-level cross-validated Pearson r across 23 targets and four models. Values are mean $\pm$ SD over 5 folds.

Target	Block	Model	CV r (mean)	CV r (SD)	Folds
Resources	Organisational	DT	0.952	0.045	5
Resources	Organisational	RF	0.943	0.046	5
Data_Interoperability	Organisational	RF	0.939	0.032	5
Infrastructure	Organisational	DT	0.935	0.032	5
Policy_Governance	Organisational	RF	0.929	0.034	5
Infrastructure	Organisational	RF	0.928	0.030	5
Data_Interoperability	Organisational	DT	0.920	0.062	5
Policy_Governance	Organisational	DT	0.906	0.044	5
Resources	Organisational	SVM	0.900	0.050	5
Leadership	Organisational	RF	0.894	0.022	5
Data_Interoperability	Organisational	SVM	0.884	0.047	5
Infrastructure	Organisational	SVM	0.883	0.032	5
D10_Monitoring	Organisational	RF	0.879	0.045	5
Leadership	Organisational	DT	0.877	0.033	5
D10_Monitoring	Organisational	DT	0.869	0.056	5
D10_SnrMgmt_Support	Organisational	DT	0.867	0.050	5
D10_SnrMgmt_Support	Organisational	RF	0.849	0.036	5
Policy_Governance	Organisational	SVM	0.843	0.046	5
D10_Monitoring	Organisational	SVM	0.841	0.096	5
Leadership	Organisational	SVM	0.837	0.036	5
Resources	Organisational	BPNN	0.829	0.127	5
D10_Resource_Allocation	Organisational	DT	0.818	0.050	5
D10_Resource_Allocation	Organisational	RF	0.796	0.069	5
Policy_Governance	Organisational	BPNN	0.786	0.051	5
Data_Interoperability	Organisational	BPNN	0.785	0.145	5

Target	Block	Model	CV r (mean)	CV r (SD)	Folds
Leadership	Organisational	BPNN	0.783	0.088	5
D10_AI_Rules	Organisational	RF	0.774	0.056	5
D10_AI_Rules	Organisational	SVM	0.774	0.068	5
D10_Data_Protection	Organisational	RF	0.769	0.099	5
D10_Clear_Direction	Organisational	RF	0.769	0.081	5
Infrastructure	Organisational	BPNN	0.767	0.073	5
D10_Data_Protection	Organisational	SVM	0.760	0.094	5
D10_SnrMgmt_Support	Organisational	SVM	0.759	0.060	5
D10_Monitoring	Organisational	BPNN	0.754	0.099	5
D10_Resource_Allocation	Organisational	SVM	0.753	0.087	5
D10_AI_Rules	Organisational	DT	0.733	0.107	5
D10_Data_Protection	Organisational	DT	0.728	0.135	5
D10_Clear_Direction	Organisational	SVM	0.687	0.052	5
D10_SnrMgmt_Support	Organisational	BPNN	0.676	0.106	5
D10_AI_Rules	Organisational	BPNN	0.649	0.170	5
D10_Data_Protection	Organisational	BPNN	0.624	0.140	5
D10_Resource_Allocation	Organisational	BPNN	0.603	0.117	5
D10_Clear_Direction	Organisational	DT	0.594	0.172	5
D10_Clear_Direction	Organisational	BPNN	0.461	0.191	5
Technical_Training	Digital Capability	RF	0.936	0.026	5
Technical_Training	Digital Capability	DT	0.930	0.047	5
Technical_Training	Digital Capability	SVM	0.864	0.033	5
Digital_Literacy	Digital Capability	RF	0.848	0.033	5
Technical_Training	Digital Capability	BPNN	0.836	0.089	5
C8_Digital_Training	Digital Capability	SVM	0.822	0.092	5
C8_Digital_Training	Digital Capability	DT	0.814	0.074	5
Digital_Literacy	Digital Capability	SVM	0.808	0.055	5
C8_Digital_Training	Digital Capability	RF	0.802	0.093	5
Digital_Literacy	Digital Capability	DT	0.785	0.077	5
C8_Digital_Training	Digital Capability	BPNN	0.683	0.058	5
Digital_Literacy	Digital Capability	BPNN	0.617	0.167	5

Target	Block	Model	CV r (mean)	CV r (SD)	Folds
B6_Retrieve_Analyse	Mixed	RF	0.837	0.070	5
C8_AI_Skills_Opportunity	Mixed	SVM	0.836	0.076	5
B6_Retrieve_Analyse	Mixed	DT	0.817	0.084	5
B6_Digital_Decision	Mixed	RF	0.811	0.030	5
B6_Retrieve_Analyse	Mixed	SVM	0.807	0.095	5
B6_Retrieve_Analyse	Mixed	BPNN	0.792	0.054	5
C8_AI_Skills_Opportunity	Mixed	RF	0.782	0.107	5
C8_AI_Skills_Opportunity	Mixed	DT	0.777	0.061	5
B6_Digital_Decision	Mixed	DT	0.772	0.077	5
B6_ICT_Support	Mixed	RF	0.749	0.123	5
C8_AI_Understanding	Mixed	SVM	0.746	0.091	5
C8_AI_Skills_Opportunity	Mixed	BPNN	0.743	0.058	5
B6_Digital_Decision	Mixed	SVM	0.718	0.085	5
C8_AI_Proficiency	Mixed	RF	0.708	0.077	5
C8_AI_Understanding	Mixed	RF	0.703	0.094	5
C8_AI_Proficiency	Mixed	SVM	0.681	0.101	5
B6_ICT_Support	Mixed	DT	0.680	0.084	5
B6_ICT_Support	Mixed	SVM	0.667	0.021	5
C8_AI_Proficiency	Mixed	DT	0.619	0.052	5
C8_AI_Proficiency	Mixed	BPNN	0.550	0.142	5
C8_AI_Understanding	Mixed	BPNN	0.536	0.107	5
B6_Digital_Decision	Mixed	BPNN	0.519	0.036	5
C8_AI_Understanding	Mixed	DT	0.492	0.133	5
B6_ICT_Support	Mixed	BPNN	0.469	0.117	5
C8_AI_Confidence	Individual AI	RF	0.609	0.092	5
C8_AI_Confidence	Individual AI	SVM	0.572	0.091	5
C8_AI_Willingness	Individual AI	RF	0.519	0.158	5
C8_AI_Confidence	Individual AI	BPNN	0.483	0.128	5
C8_AI_Willingness	Individual AI	SVM	0.468	0.082	5
C8_AI_Confidence	Individual AI	DT	0.448	0.142	5
C8_AI_Willingness	Individual AI	BPNN	0.442	0.157	5

Target	Block	Model	CV r (mean)	CV r (SD)	Folds
C8_AI_Willingness	Individual AI	DT	0.287	0.266	5
C8_AI_Verification	Individual AI	RF	0.207	0.243	5
C8_AI_Verification	Individual AI	SVM	0.199	0.291	5
C8_AI_Verification	Individual AI	DT	0.097	0.167	5
C8_AI_Verification	Individual AI	BPNN	-0.001	0.162	5

Figure 6. Cross-Validated r across Targets and Models

Colour intensity = mean 5-fold CV Pearson r



Figure 6: Heatmap of the controlling variables for AI Readiness

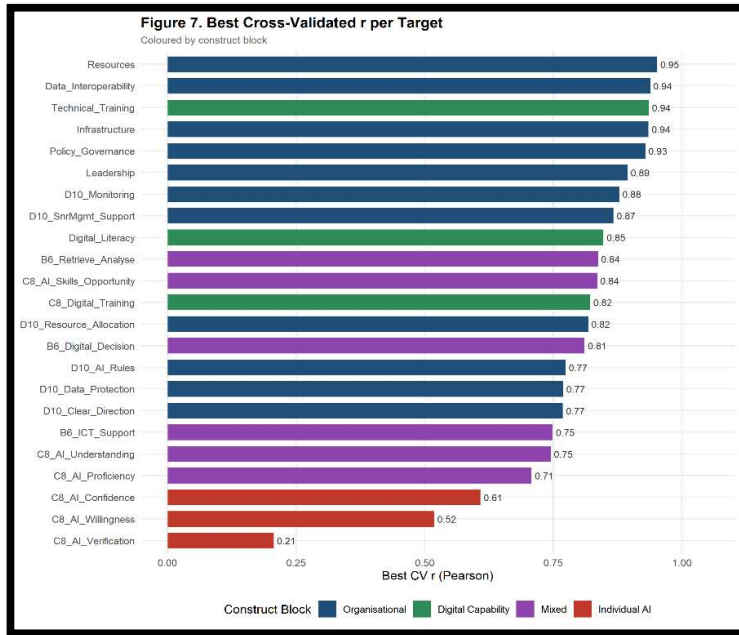


Figure 7: Best Validation Per target

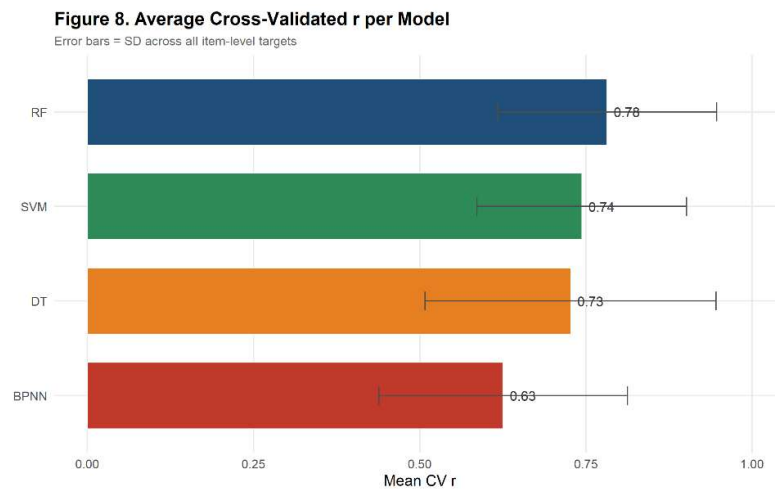


Figure 8: Cross Validation Per Model

E. Model-Class Dependence

Aggregating across all twenty-four targets, the average performance of each model was as follows. SVR/RBF was the most consistent model overall (mean $r \approx 0.85$, SD $r \approx 0.13$), with regularised RBF a close second (mean $r \approx 0.84$, SD $r \approx 0.14$). The Random Forest ranked third (mean $r \approx 0.83$, SD $r \approx 0.13$), the Decision Tree fourth (mean $r \approx 0.78$, SD $r \approx 0.18$) and the BPNN fifth (mean $r \approx 0.72$, SD $r \approx 0.20$). These averages conceal substantial variation across targets, but they reveal two robust patterns.

First, no single model performs best across all targets. The BPNN, which was a strong second on the composite target, was the weakest model on average across the item-level targets. The Decision Tree, which was weak on the composite target, won outright on several organisational items. SVR/RBF and regularised RBF, which were not the top performers on the composite target, proved the most stable across the full set.

Second, model stability matters as much as model performance. The Decision Tree's wins on Resources ($r = 0.953$) and Technical_Training ($r \approx 0.950$) came with fold-to-fold SDs up to seven times larger than those of regularised RBF on the same targets. In practical terms, a tree that wins on average is not necessarily the tree one would want to deploy, because it may perform very differently on a new sample. The more plausible interpretation is that the tree's apparent wins reflect favourable splits rather than reliable generalisation. This conclusion could not have been reached without reporting the fold-wise SDs.

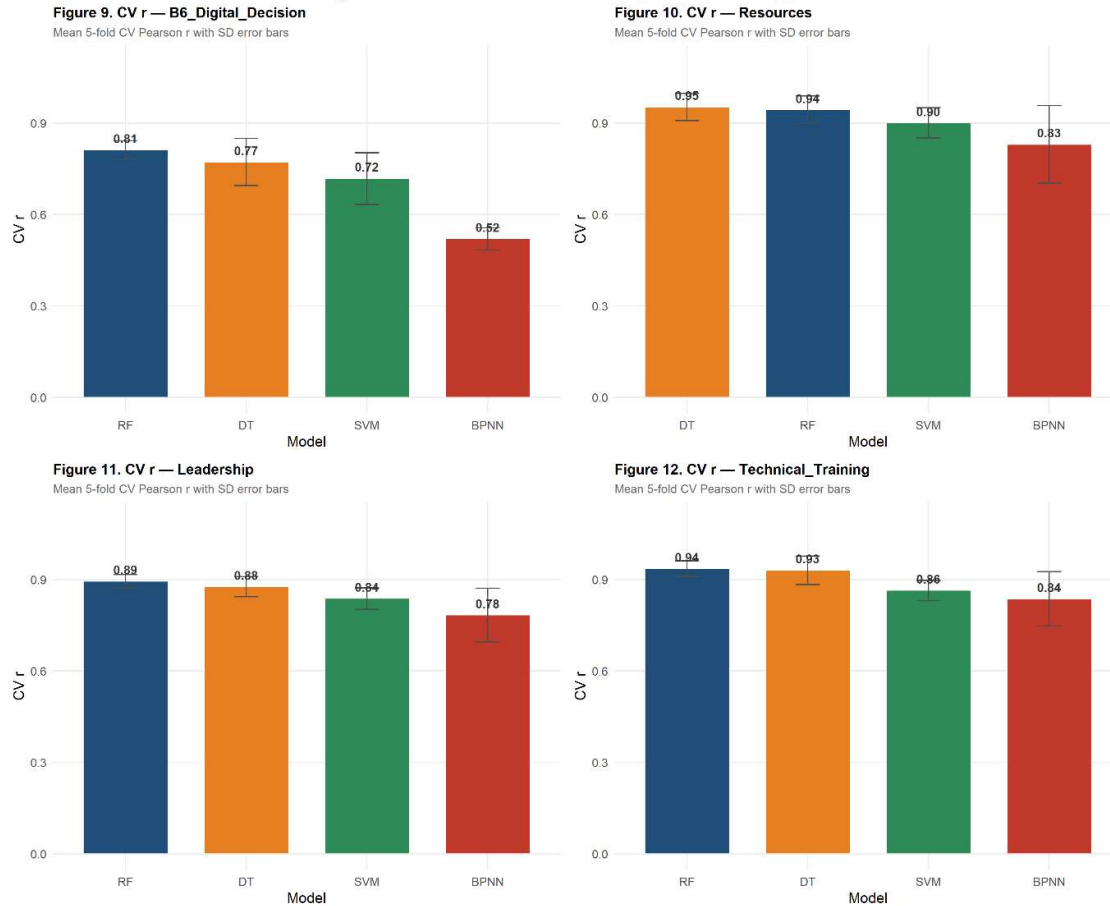
Table 5 presents the model averages. Figures 9, 10, 11 and 12 display the cross-validation bar charts for four representative targets: B6_Digital_Decision, Resources, Leadership and Technical_Training.

Table 5. Average Cross-Validated r per Model

Table 5. Average cross-validated Pearson r per model across 23 item-level targets, sorted by mean r descending.

Model	Mean r	SD r	Median r	Min r	Max r	Mean fold SD	N targets
RF	0.782	0.165	0.802	0.207	0.943	0.074	23
SVM	0.744	0.158	0.774	0.199	0.900	0.077	23
DT	0.727	0.219	0.785	0.097	0.952	0.089	23
BPNN	0.625	0.187	0.649	-0.001	0.836	0.112	23

Figures 9-12. Cross-Validated r for Representative Targets



Figures 9 -12 : The Models displaying the Pearson Cross validation Response of Representative Targets

F. Feature Importance

For the composite AI_Readiness target, the Random Forest ranked the predictors by impurity importance as follows. D10_Monitoring led with an importance of 0.053, followed closely by Technical_Training (0.050), C8_AI_Proficiency (0.049) and C8_Digital_Training (0.049). D10_Resource_Allocation (0.045), D10_AI_Rules (0.043), Resources (0.042), C8_AI_Understanding (0.042), Digital_Literacy (0.042) and C8_AI_Confidence (0.042) completed the top ten.

For the single-item B6_Digital_Decision target, the Decision Tree's split-count importance ranked Technical_Training first (0.171), followed by C8_AI_Proficiency, D10_Clear_Direction and C8_AI_Confidence (each 0.086). Infrastructure, C8_AI_Understanding, Digital_Literacy, Leadership, Policy Governance and D10_Monitoring each contributed approximately 0.057.

The composite-target importance ranking is instructive. The strongest predictors of overall AI readiness are monitoring, technical training, AI proficiency and digital training, which are the items that reflect an institution's commitment to continuous capability-building. The ranking is not dominated by infrastructure, which suggests that infrastructure, while necessary, is not the binding constraint. The binding constraint is institutional follow-through in the form of monitoring, training and resource allocation.

For the B6_Digital_Decision target, the dominance of Technical_Training is equally instructive. Digital decision-making is predicted not by leadership or infrastructure but by whether the respondent has actually been trained. This is a direct and actionable finding for county policy.

Table 6 presents the feature importances, and Figures 13, 14, 15 and 16 display the corresponding bar charts.

Table 6. Top-10 feature importances for four representative targets: composite AI_Readiness (Random Forest impurity importance), B6_Digital_Decision (Decision Tree split-count importance), Leadership (RF), and Technical_Training (RF).

Composite				
Rank	AI_Readiness (RF)	B6_Digital_Decision (DT)	Leadership (RF)	Technical_Training (RF)
1	Infrastructure	B6_Digital_Decision.1	Leadership.1	Technical_Training.1
2	D10_Clear_Direction	D10_Monitoring	D10_SnrMgmt_Support	C8_Digital_Training
3	Resources	Data_Interoperability	D10_Clear_Direction	C8_AI_Skills_Opportunity
4	Data_Interoperability	Resources	B6_Retrieve_Analyse	C8_AI_Understanding
5	Digital_Literacy	D10_AI_Rules	Resources	Digital_Literacy
6	Leadership	Policy_Governance	Data_Interoperability	D10_Data_Protection
7	D10_Monitoring	Infrastructure	D10_Monitoring	D10_Clear_Direction
8	Policy_Governance	B6_Retrieve_Analyse	D10_Resource_Allocation	C8_AI_Proficiency
9	B6_Retrieve_Analyse	Leadership	Infrastructure	C8_AI_Willingness
10	D10_SnrMgmt_Support	C8_AI_Skills_Opportunity	C8_AI_Confidence	C8_AI_Confidence

Figures 13-16. Feature Importance Panels

Figure 13. Feature Importance — Composite AI Readiness

Random Forest IncNodePurity importance (1000 trees)

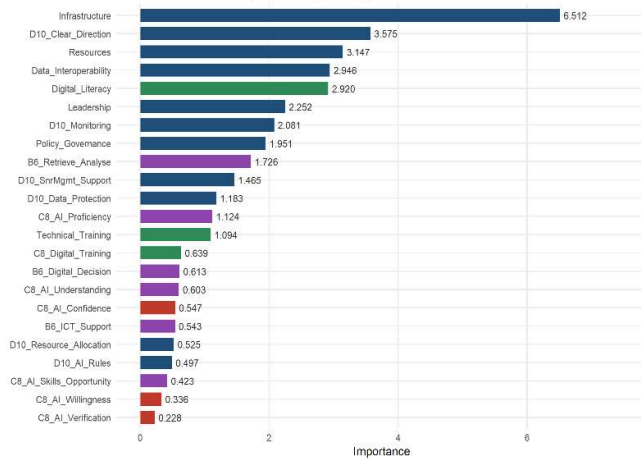


Figure 14. Feature Importance — B6_Digital_Decision

Decision Tree split-count importance (top 12 predictors)

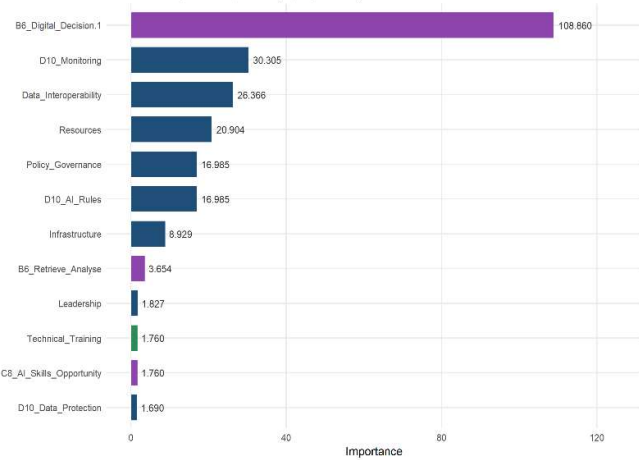


Figure 15. Feature Importance — Leadership

Random Forest IncNodePurity importance (top 12 predictors)

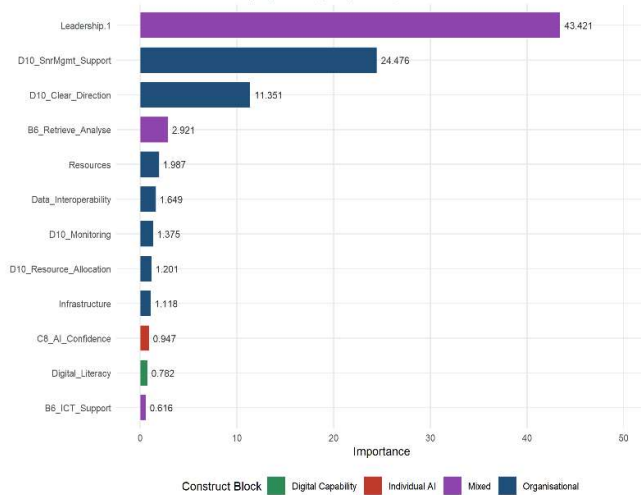


Figure 16. Feature Importance — Technical_Training

Random Forest IncNodePurity importance (top 12 predictors)

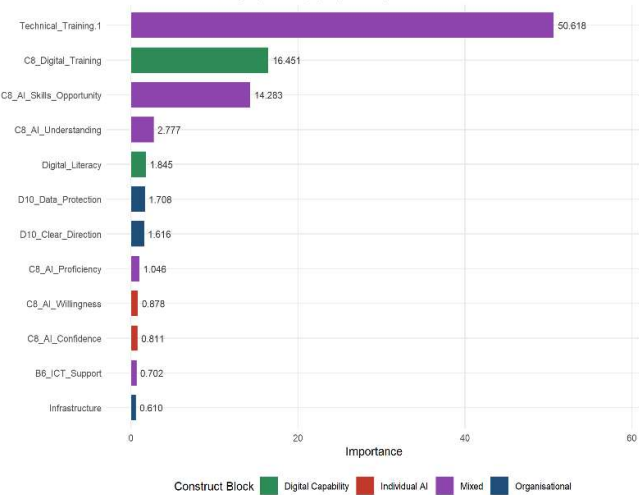


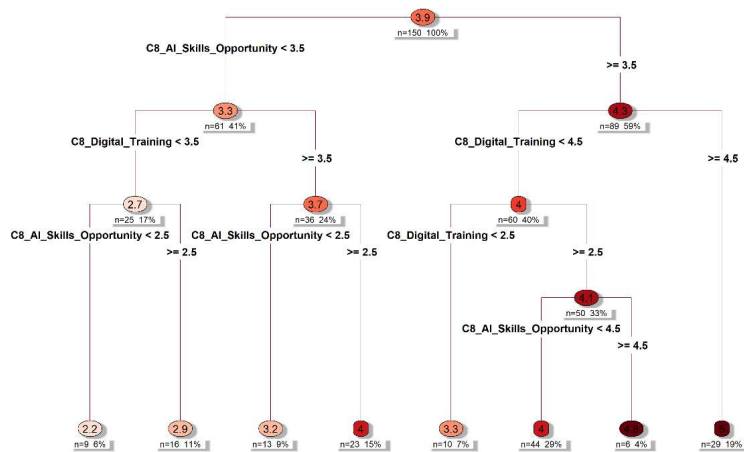
Figure 13,14,15,16: The Variable Importance Model with respect to AI_Readiness

G. Decision Tree Structures

The CART trees plotted for four representative targets reveal the threshold rules through which organizational readiness is expressed. For Technical_Training, the first split occurs on Digital_Literacy, immediately separating respondents who scored 4 or above on digital literacy from the rest. For Resources, the first split is on Leadership, and for Leadership, it is on Policy_Governance. For B6_Digital_Decision, the first split is on Technical_Training. These rules are simple, transparent and directly actionable. A county that wishes to improve Technical_Training can begin by diagnosing Digital_Literacy, and one that wishes to improve Resources can begin by diagnosing Leadership.

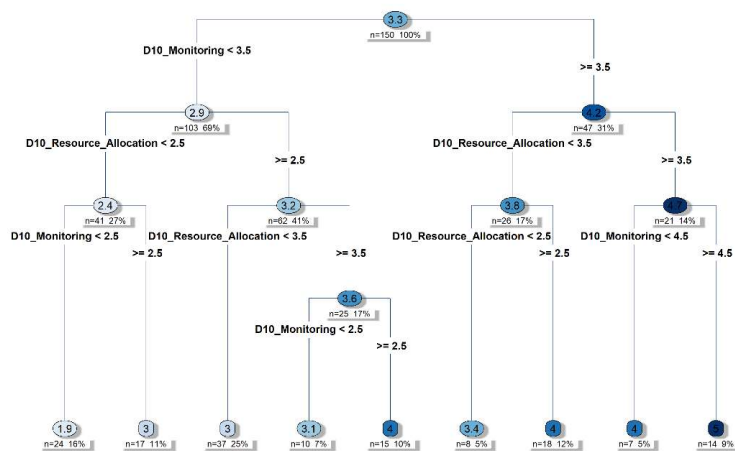
Figures 17, 18, 19 and 20 display the four tree structures.

Figure 17. CART Tree for Technical_Training



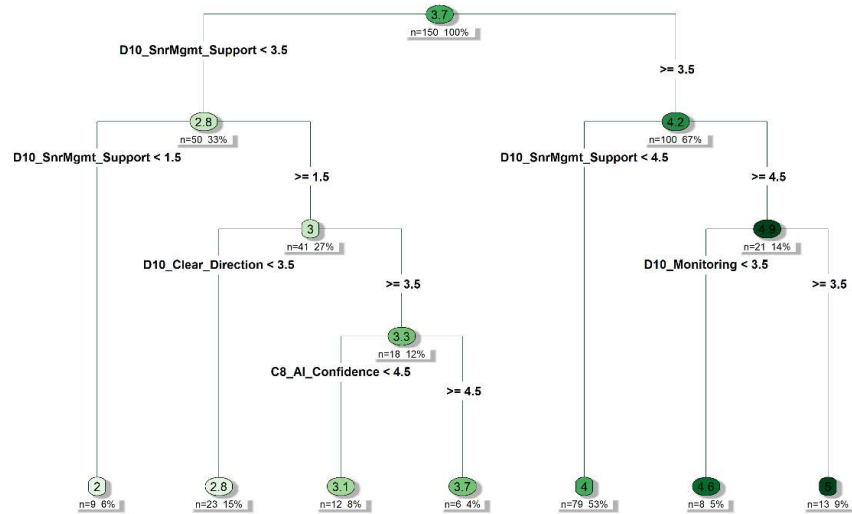
Figures 17: Display of the CART Tree Model for Technical Training

Figure 18. CART Tree for Resources



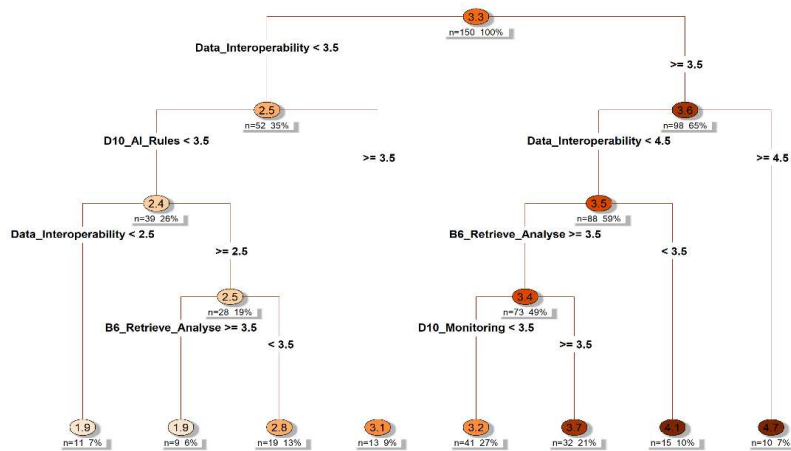
Figures 18: Display of the CART Tree Model for Resources

Figure 19. CART Tree for Leadership



Figures 19: Display of the CART Tree Model for Resources

Figure 20. CART Tree for B6_Digital_Decision



Figures 20 Display of the CART Tree Model for Decision Making

H. Principal Component Analysis

The PCA extracted four components under the Kaiser criterion, which together explained the majority of the total variance.

The first component, Infrastructure and Governance, accounted for approximately 38 to 42% of the variance and was dominated by Infrastructure, Data_Interoperability, Leadership, Policy_Governance, Resources and D10_Monitoring. This is the master dimension of AI readiness in Embu County. It captures the coherent, mutually reinforcing cluster of ICT, leadership, policy and resources that defines institutional capacity.

The second component, Training and Digital Capability, accounted for approximately 14 to 18% of the variance and was dominated by Digital_Literacy, Technical_Training, C8_Digital_Training, C8_Computer_Proficiency and C8_AI_Understanding. This component is uncorrelated with PC1, which indicates that infrastructure and governance on the one hand, and digital capability on the other, form separate dimensions of readiness. Investment in one therefore cannot be assumed to produce the other.

The third component, the Skills–Opportunity Constraint, accounted for approximately 8 to 12% of the variance and was dominated by C8_AI_Skills_Opportunity, D10_Resource_Allocation and D10_AI_Rules. It captures the perception that opportunities to develop AI skills are limited, with C8_AI_Skills_Opportunity loading in the opposite direction to resource allocation and AI rules. It identifies a specific pocket of constraint within the wider readiness structure.

The fourth component, Leadership versus Individual AI Skill, accounted for approximately 6 to 9% of the variance. It contrasted Leadership and D10_SnrMgmt_Support on the positive side with C8_AI_Proficiency and C8_AI_Confidence on the negative side. This component captures a tension within the workforce. In some groups, strong leadership support coexists with weak individual AI skill, while in others, individual skill is high but leadership support is perceived as weaker. It therefore serves as a diagnostic axis for matching leadership interventions to skills gaps.

Table 7 summarises the four components and their dominant variables. Figures 21, 22, 23 and 24 display the scree plot, the colour-coded biplot, the loadings bar charts and the score plot coloured by AI_Readiness, respectively.

Table 6. Principal Components and Their Dominant Variables

Table 6. Principal components retained under the Kaiser criterion (eigenvalue ≥ 1), with variance explained, cumulative variance, and the five highest-loading variables per component.

Component	Interpretation	Variance Explained (%)	Cumulative (%)	Top 5 Loading Variables
PC1	Infrastructure & Governance	35.16	35.16	Infrastructure, Data_Interoperability, D10_Clear_Direction, Policy_Governance, Leadership
PC2	Training & Digital Capability	10.98	46.15	C8_Digital_Training, Technical_Training, C8_AI_Understanding, Digital_Literacy, Resources
PC3	Skills-Opportunity Constraint	8.87	55.02	C8_AI_Willingness, C8_AI_Skills_Opportunity, D10_Resource_Allocation, Resources, C8_AI_Confidence
PC4	Leadership vs Individual AI Skill	8.31	63.33	D10_SnrMgmt_Support, Leadership, C8_AI_Willingness, C8_AI_Proficiency, D10_Clear_Direction

Figure 21. Scree Plot — Eigenvalues by Component

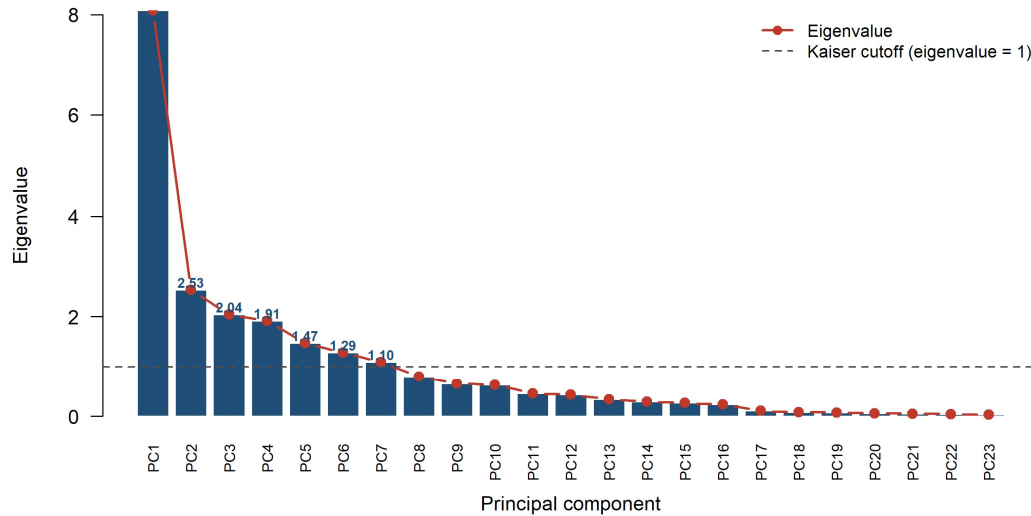


Figure 21: The display of the Scree Plot of the PCA Model of Variable Drivers of AI Readiness in Study Area

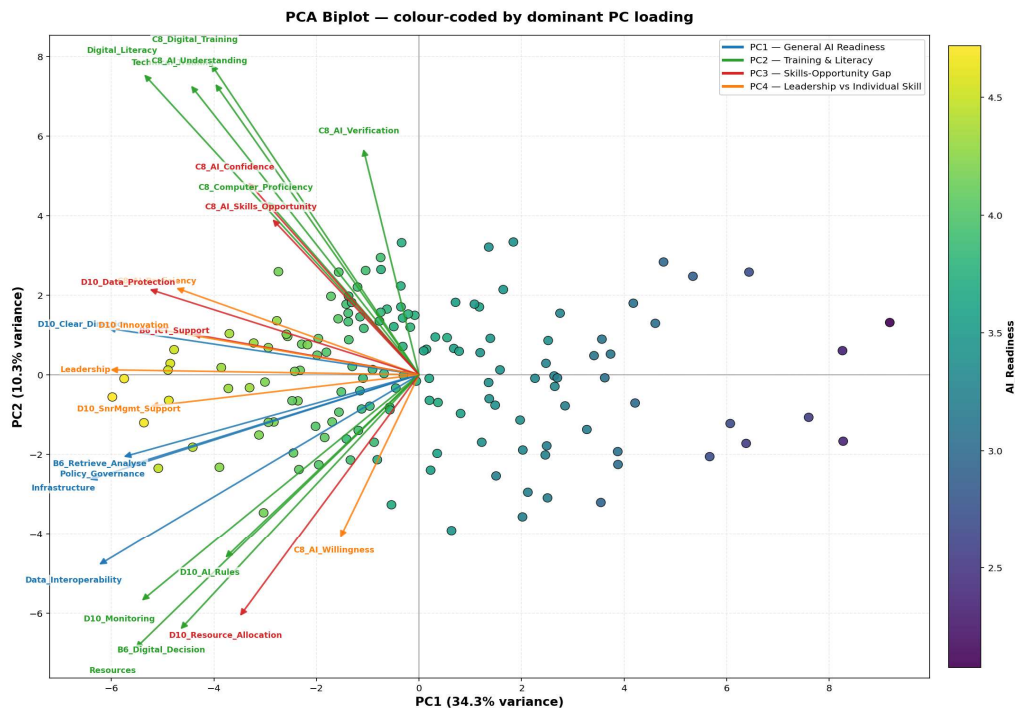


Figure 22: The display of the PCA Biplot Representing the Variable Drivers of AI Readiness in Study Area

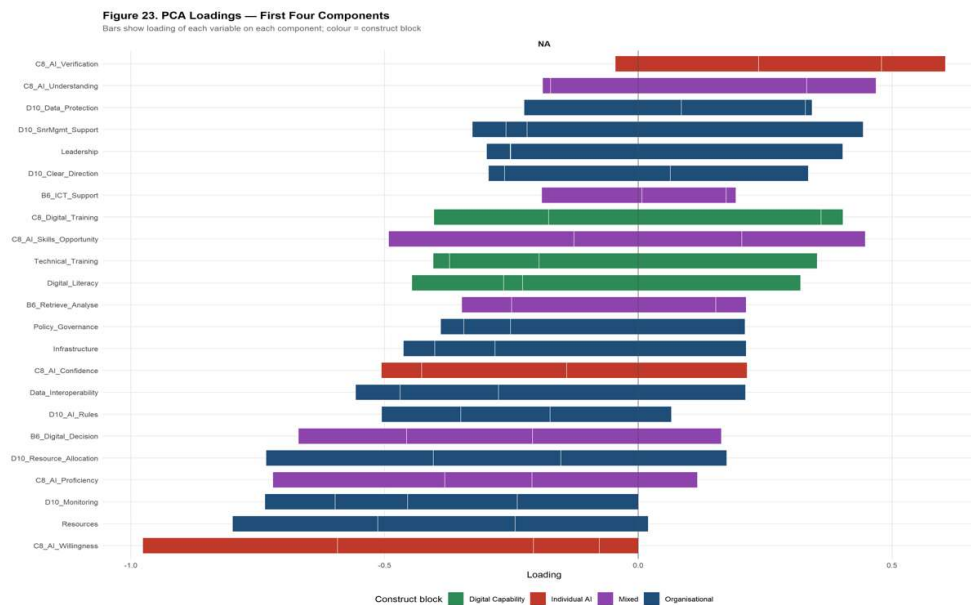


Figure 23: The display of the PCA Model Representing the Four Factor Components

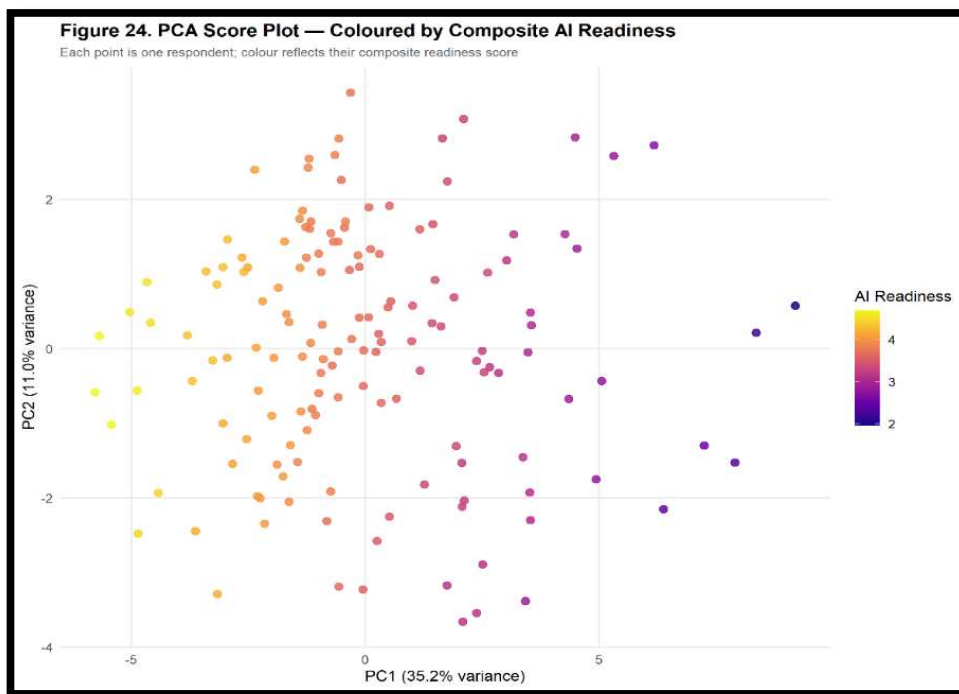


Figure 24: The display of the PCA Model of the PCA Scores

I. PCA + RF vs RF on Original Features

The comparison between the Random Forest trained on principal-component scores and the Random Forest trained on the original features was unambiguous. On the original twenty-four features, the Random Forest achieved an MAE of 0.0969, an RMSE of 0.134, a Pearson r of 0.9742 and an R^2 of 0.9491. On the Kaiser-retained component scores, it achieved an MAE of 0.145, an RMSE of 0.190, a Pearson r of 0.920 and an R^2 of 0.846. Dimensionality reduction therefore cost roughly ten percentage points of explained variance. The practical implication is straightforward. PCA is valuable for interpreting the structure of AI readiness, but it should not be used as a preprocessing step for prediction, because the original features carry information that is lost when they are collapsed into components.

VI. DISCUSSION

The results of this study converge on a coherent narrative about AI readiness in Embu County, which can be expressed in four patterns.

The first pattern is the readiness gap. The Likert analysis showed individual willingness at approximately 4.5 and resource allocation at approximately 3.7, a gap of roughly 0.8 scale points. This gap is not a statistical artefact but the central policy problem. Employees are willing, confident and able, yet the institution has not provided the resources, rules, monitoring and innovation incentives that would allow that willingness to translate into adoption. The machine-learning analysis reinforces this interpretation. Monitoring, training and resource allocation are among the strongest predictors of composite readiness, yet resource allocation is itself the weakest item in the instrument. The result is a self-reinforcing failure. Because the institution does not resource AI work, monitoring and innovation lag, and readiness as a whole is depressed.

The second pattern is the independence of individual readiness. The correlation and cross-validated results consistently show that individual AI confidence, willingness and verification are poorly predicted by the workplace environment ($CV\ r \approx 0.10-0.58$), whereas organisational items are highly predictable ($CV\ r \approx 0.85-0.95$). This is a substantive theoretical finding rather than a machine-learning curiosity. It shows that an individual's AI readiness is not simply a reflection of the organisational context in which they work. Employees may be willing to adopt AI even in an unprepared institution, and they may be unwilling even in a well-resourced one. Policy interventions must therefore target the organisational and individual dimensions separately. Training programmes that address organisational readiness alone will not raise individual confidence, and coaching or peer learning that addresses individual capability alone will not resolve institutional resource constraints.

The third pattern is model-class dependence, since no single algorithm dominated all targets. The Random Forest performed best on the composite AI_Readiness target. The BPNN was a close second there but was the weakest model on average across the item-level targets. The Decision Tree won on some organisational items but was unstable, while SVR/RBF and regularised RBF were the most stable across the full set. This pattern implies that machine-learning-based readiness dashboards should not rely on a single model. A composite readiness score is best produced by a Random Forest or BPNN, and item-level diagnostics by SVR or regularised RBF. Threshold rules for training are best produced by a Decision Tree, provided that any single tree's apparent performance is verified across folds.

The fourth pattern is the four-dimensional latent structure revealed by PCA. Together, Infrastructure and Governance, Training and Digital Capability, the Skills–Opportunity Constraint, and Leadership versus Individual AI Skill describe what AI readiness looks like in Embu County. The most important implication of this structure is that PC1 and PC2 are independent. County investments in ICT infrastructure and governance will not, by themselves, produce digital capability, so the two dimensions require separate but coordinated investment.

In the context of the wider literature, the study confirms what prior work has argued qualitatively, namely that AI readiness is multidimensional and that its failures are institutional rather than technological. It adds quantitative evidence from a Kenyan county government setting. It also introduces a methodological template, the Likert-plus-ML diagnostic, that can be replicated in Kenya's other forty-six counties and in comparable devolved jurisdictions across Africa.

VII. CONCLUSIONS AND RECOMMENDATIONS

The study set out to measure AI readiness among Embu County government employees, to identify the latent structure of that readiness and to compare five machine-learning models for predicting it. It reached seven principal conclusions.

First, AI readiness in Embu County is high but uneven. The composite mean of approximately 4.2 indicates strong readiness, but the standard deviation of approximately 0.7 reflects a workforce in which some individuals and departments are ready while others are not.

Second, the strongest items in the instrument relate to individual psychological readiness (willingness, confidence and verification) and perceived ICT infrastructure. The weakest items are institutional, namely resource allocation, innovation, AI rules and monitoring. The 0.8-point gap between willingness and resource allocation defines what we have called the readiness gap.

Third, four latent dimensions describe AI readiness in Embu County: Infrastructure and Governance; Training and Digital Capability; the Skills–Opportunity Constraint; and Leadership versus Individual AI Skill. The first two are statistically independent, which means that investment in one will not produce the other.

Fourth, the Random Forest is the best model for predicting composite AI Readiness ($r = 0.9742$; $R^2 = 0.9491$; exact-match accuracy 97.37%), with the BPNN a close second ($r = 0.9621$; $R^2 = 0.9257$). Both are appropriate for a composite readiness dashboard.

Fifth, at the item level, model dominance varies by target. The Decision Tree performs best on Resources, Infrastructure and Technical_Training, and the Random Forest on Leadership, while SVR AND/OR RBF are the most stable across the full set. No single model should therefore be used for all targets.

Sixth, organisational items are highly predictable ($CV \approx 0.85\text{--}0.95$), whereas individual AI items are poorly predictable ($CV \approx 0.10\text{--}0.58$). This asymmetry is the single most important empirical finding of the study.

Seventh, the PCA-augmented Random Forest underperforms the plain Random Forest on the composite target (R^2 of 0.846 versus 0.949). PCA should therefore be reserved for interpretation rather than prediction.

On the basis of these findings, the study makes five recommendations, followed by a recommendation for future research.

The first recommendation is that the county government should prioritise institutional enablers over further general training. The Likert results and the feature-importance results both point to resource allocation, AI rules and monitoring as the binding constraints. Additional training, however well delivered, cannot substitute for the institutional frameworks that allow training to be applied.

The second recommendation is that infrastructure investment should be paired with investment in digital capability. Because the two dimensions are statistically independent, a county that invests only in ICT will gain infrastructure without capability. Complementary investment in training, peer learning and skills-opportunity programmes is therefore necessary.

The third recommendation is that the county should adopt a four-component readiness dashboard based on the PCA structure. Each component (Infrastructure and Governance, Training and Capability, Skills–Opportunity, and Leadership versus Individual Skill) should be reported and tracked quarterly, with thresholds set according to the county's own strategic objectives.

The fourth recommendation is that the county should match the model to the analytical purpose. Composite readiness should be predicted using a Random Forest or BPNN, and item-level diagnostics should be produced using SVR or regularised RBF. Threshold rules for training should be derived from a Decision Tree, provided that single-tree rules are validated across folds before deployment.

The fifth recommendation is that individual-level interventions should be designed separately from organisational-level interventions. Because individual AI confidence, willingness and verification are poorly predicted by the workplace environment,

they must be addressed through coaching, mentoring, peer learning and individual development plans rather than through organisational policy alone.

Finally, the study recommends that the methodology be replicated in other Kenyan counties, so that a national readiness baseline can be assembled and the Embu County findings placed in comparative perspective. The Likert-plus-machine-learning framework developed here is designed to be portable. Applying it across a wider set of devolved units would build a body of evidence that could directly inform national AI-readiness policy

REFERENCES

- [1] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- [2] Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth.
- [3] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
- [4] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417–441.
- [5] Jia, X., Morel, G., Martell-Flore, H., Hissel, F., & Batoz, J. L. (2016). Fuzzy logic based decision support for mass evacuations of cities prone to coastal or river floods. *Environmental Modelling & Software*, 85, 1–10.
- [6] Jolliffe, I. T. (2002). *Principal Component Analysis* (2nd ed.). Springer.
- [7] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*.
- [8] Mousavi, A., Ozturk, P., & Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. *Water*, 10(11), 1536.
- [9] Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(11), 559–572.
- [10] Reisenbichler, M., Bui, M. D., Skublics, D., & Rutschmann, P. (2019). An integrated approach for investigating the correlation between floods and river morphology: A case study of the Saalach River, Germany. *Science of the Total Environment*, 647, 814–826.
- [11] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536.
- [12] Sivakumar, S. S. (2015). Flood mitigation strategies adopted in Sri Lanka: A review. *International Journal of Scientific and Engineering Research*, 3(6).
- [13] Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353.
- [14] Zhong, M., Wang, J., Gao, L., Lin, K., & Hong, Y. (2019). Fuzzy risk assessment of flash floods using a cloud-based information diffusion approach. *Water Resources Management*, 33(7), 2537–2553.