SITE 2025 36th INTERNATIONAL CONFERENCE
SOCIETY FOR INFORMATION TECHNOLOGY AND TEACHER EDUCATION

Predicting Computational Thinking (CT) Skill Level Among Engineering Students Using Machine Learning (ML)

Presentation by Opeyemi Ojajuni

Co-Authors: Yasser Ismail, Brain Warren, Fareed Dawan, Albertha Lawson

National Science Foundation

NSF-funded research (Award No. 1915520)

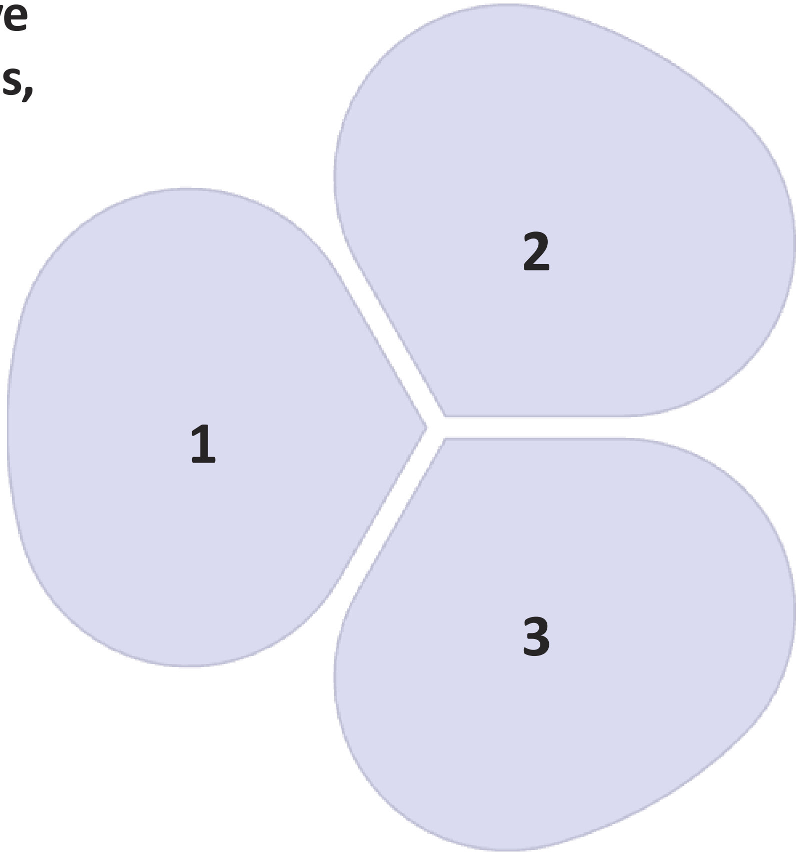
Background

Essential 21st Century Skill

- **Computational Thinking (CT)** involves cognitive cognitive abilities like breaking down problems, problems, recognizing patterns, and designing designing algorithms.

- CT essential for Artificial Intelligence (AI) and and engineering careers.

- CT assessment includes evaluating skills in **problem-solving, algorithmic thinking, critical critical thinking, cooperative thinking, and creativity**



Industry Demand

Today's technical challenges require professionals who can decompose problems, recognize patterns, and develop algorithmic solutions—core CT competencies.

Growing AI Demand

In 2023, the World Economic Forum predicts that 75% of companies will adopt AI by 2027, further driving demand for professionals skilled in CT.

Artificial Intelligence(AI): Building machines that can perform tasks such as perception, reasoning, learning, and problem-solving.

The main aspects of AI are Machine Learning (ML), Deep Learning (DL) Natural Language Processing (NLP - GenAI) , and Computer Vision.

(Brennan & Resnick, 2012; Korkmaz et al., 2017; O. Ojajuni, Ismail, Mellieon-Williams, et al., 2023; The Future of Jobs Report 2023, n.d.)

Gap in Literature/Problem Statements

1

Assessment Challenges - Self-Report Limitations

Current assessment approaches rely heavily on self-reporting measures and measures and standardized testing, which may not accurately capture the capture the multidimensional nature of computational thinking skills as they skills as they manifest in authentic engineering contexts.

2

Limited AI Integration

The literature reveals a significant shortage of research utilizing machine learning to predict computational thinking proficiency specifically among engineering students, despite the growing importance of these skills in engineering practice.

3

Curriculum Integration

Few studies have explored how insights derived from machine machine learning models can be systematically incorporated into incorporated into engineering curriculum design to enhance CT CT instruction and create personalized learning pathways.

Objectives

1

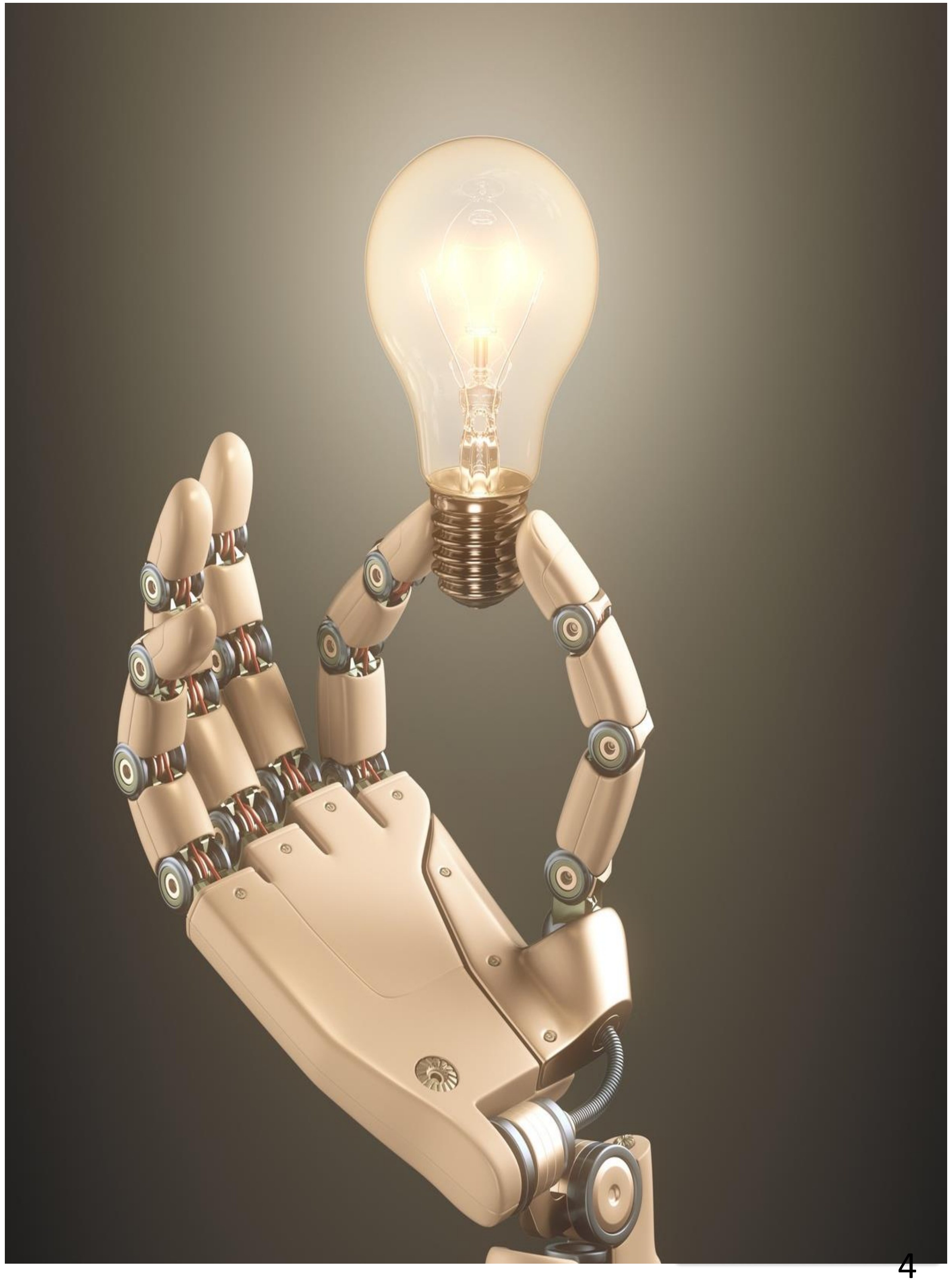
Develop Predictive Machine Learning Models

2

Identify key Factors of Computational Thinking

3

Enhance Curriculum



Research Questions

RQ1: Key Factors

What are the main factors impacting computational thinking skills among engineering students?

This question seeks to identify the critical factors that contribute most significantly to CT proficiency, providing a foundation for targeted educational interventions.

RQ2: Prediction Accuracy

How accurately can machine learning models predict computational thinking skill levels levels using survey data?

This question examines whether ML whether ML approaches can serve serve as reliable assessment tools, tools, potentially reducing the resource burden of traditional evaluation methods.

RQ3: Model Comparison

Which machine learning model achieves the highest prediction accuracy for CT accuracy for CT proficiency?

By comparing multiple ML algorithms, we aim to identify the most effective effective approach for CT skill assessment in engineering education contexts. contexts.

Methodology

Dataset

- Primary data collected from 415 engineering students at an HBCU.
- Survey instrument- Computational Thinking (CT) Scales, published by Computers in Human Behavior (2017) by Korkmaz et al,
- Survey instrument with 29 observed variables grouped into 5 CT factors: *Creative Thinking (CR), Algorithmic Thinking (AT), Cooperative Thinking (CO), Critical Thinking (CRT), Problem-Solving (PS)*
- 5- Likert scale rating: 1 (Never) to 5 (Always).

- Creativity Thinking: Ability to think innovatively and generate original ideas.
- Algorithmic Thinking: Ability to conceptualize, design, and implement algorithms or step-by-step procedures.
- Critical Thinking: Skill to objectively evaluate and analyze information, scenarios, or problems.
- Cooperativity: Ability to work effectively with others.
- Problem-Solving: Capacity to analyze and resolve complex issues or challenges.

Methodology

Data Transformation

- Survey responses loaded into a Pandas DataFrame, with each student as a row.
- CT factor scores (CR, AT, CO, CRT, PS) calculated as average Likert scale values.
- Overall CT score computed by averaging all five factors.
- Students classified as "High" or "Low" CT based on the median score.
- Aggregated scores used as features (X); binary CT_Class (High/Low) as target variable (y).
- 29 variables (X) as independent variables and CT_Class as the dependent variable (y).
- StandardScaler applied to normalize features for better ML model performance.

Tools: Python, Jupyter Notebook.

Methodology

ML procedure – Prediction classification model

- **Feature Selection & Engineering** – Identify key variables and create meaningful features using Random Forest and XGBoost models .
- **Data Splitting** – Divide data into training *(80%)* and testing *(20%)* sets.
- **Model Selection** – Choose appropriate ML models *(Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest, Extreme Gradient Boosting (XGBoost)).*
- **Model Training** – Train models using the training dataset.
- **Hyperparameter Tuning** – Optimize model parameters for better performance.
- **Model Evaluation** – Assess performance using accuracy, precision, recall, and F1-score.
- **Cross-Validation** – Validate model stability across multiple data splits

Tools: Python, Jupyter Notebook.

Results

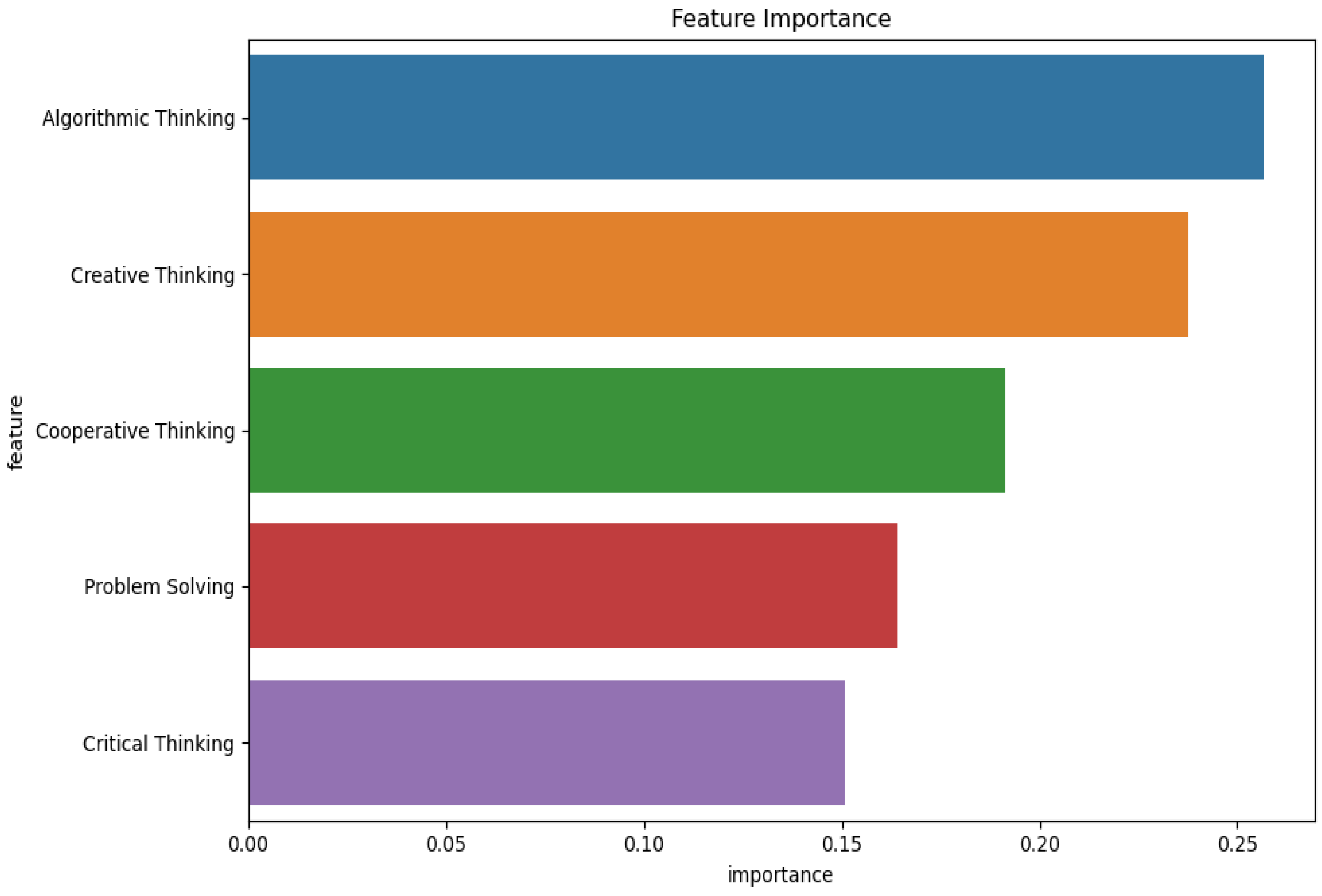


Figure 1: Factor Importance

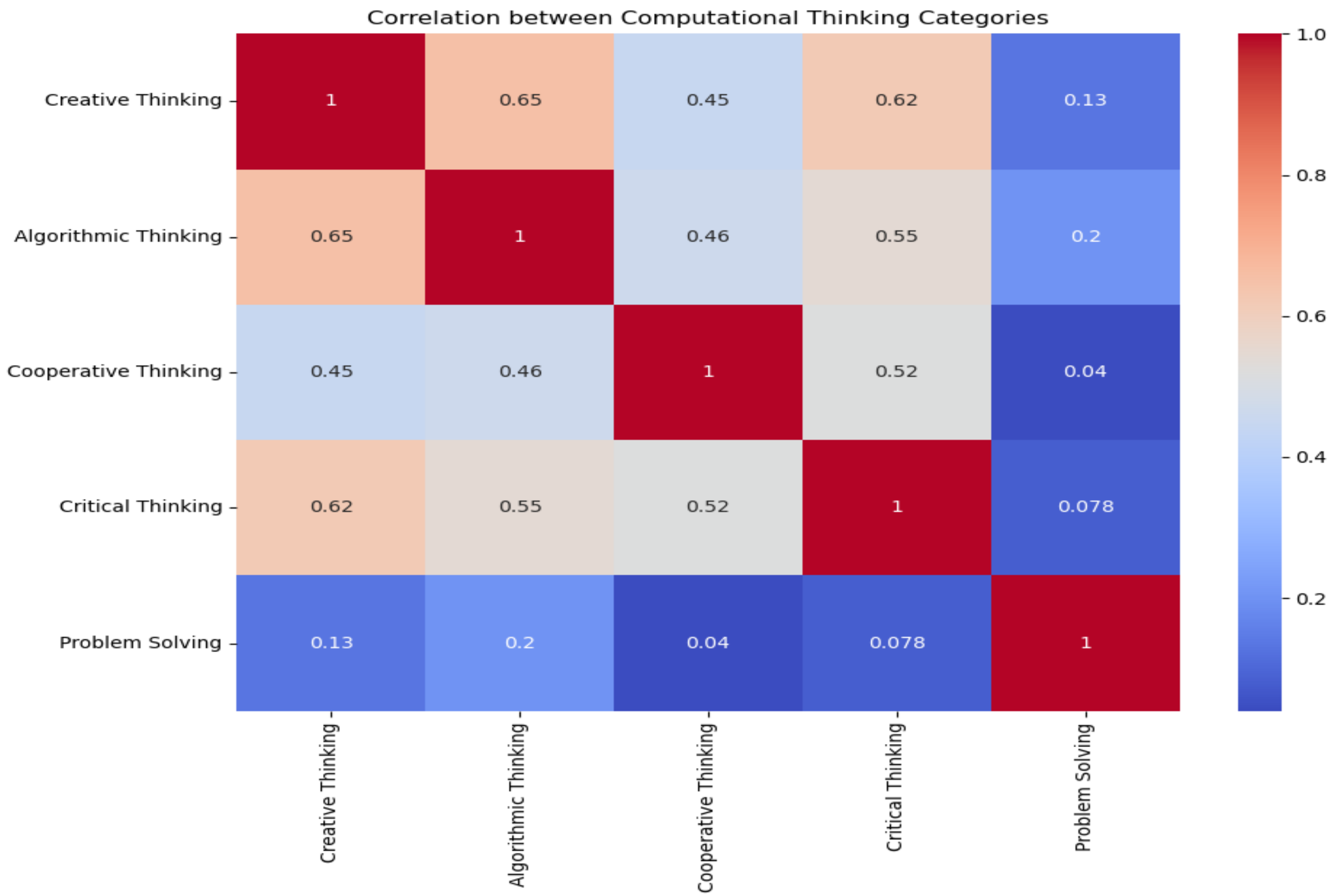


Figure 2: Correlation between computational thinking dimensions

- *Algorithmic thinking (25.67%)* **is the most influential factor, followed by** *Creative thinking (23.74%).*
- **Strong relationships:**
  - *Creative thinking & Algorithmic thinking (0.65)*
  - *Creative thinking & Critical thinking (0.62)*
- **Weak correlations for Problem-solving:**
  - *Creative thinking (0.13)*
  - *Cooperative thinking (0.04)*

Problem-solving appears to function more independently within Computational Thinking (CT).

Results

Model	Test Accuracy	CV Score	Precision (Low/High)	Recall (Low/High)	F1-Score (Low/High)
SVM	0.9518	0.9518 (±0.0555)	0.97/0.94	0.92/0.98	0.94/0.96
KNN	0.9398	0.9160 (±0.0904)	0.94/0.94	0.92/0.96	0.93/0.95
XGBoost	0.9277	0.9337 (±0.0971)	0.92/0.94	0.92/0.94	0.92/0.94
Random Forest	0.9036	0.8918 (±0.0846)	0.87/0.93	0.92/0.89	0.89/0.91

Table 1: Performance metrics of machine learning models

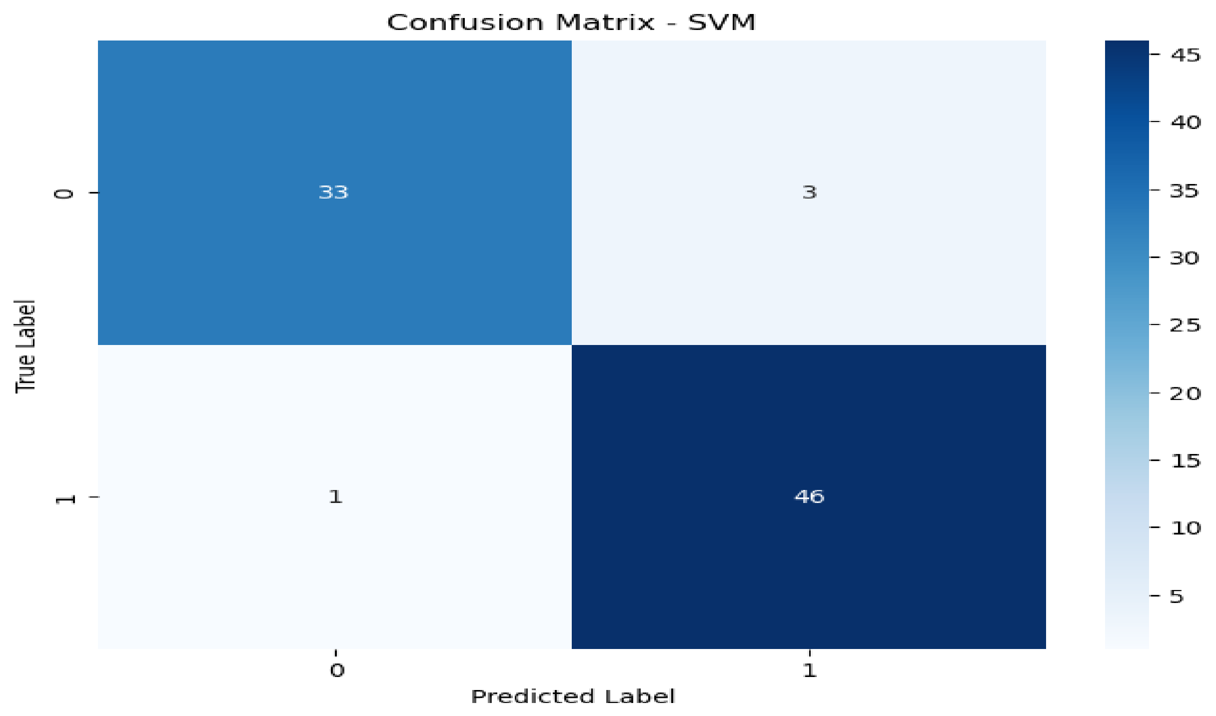


Figure 3. confusion matrix

Support Vector Machine (SVM) achieved the highest prediction accuracy at 95.18%, followed closely by K-Nearest Neighbors at 93.98%.

Neighbors at 93.98%.

- ***True Negatives (33):*** **The model correctly identified 33 instances of low CT skills.**
- ***False Positives (3):*** **Three instances of low CT skills were misclassified as high CT skills.**
- ***False Negatives (1):*** **One instance of high CT skills was misclassified as low CT skills.**
- ***True Positives (46):*** **The model correctly identified 46 instances of high CT skills.**

SVM model performs well, particularly in identifying high CT skills (with only one false negative), showing strong precision and recall.

Implication of Results

1

Machine Learning Predictive Models

- *Early identification of students needing support.*
- *Personalized learning experience.*
- *Machine Learning enables objective, scalable CT assessment.*

2

Identify Key CT Factors

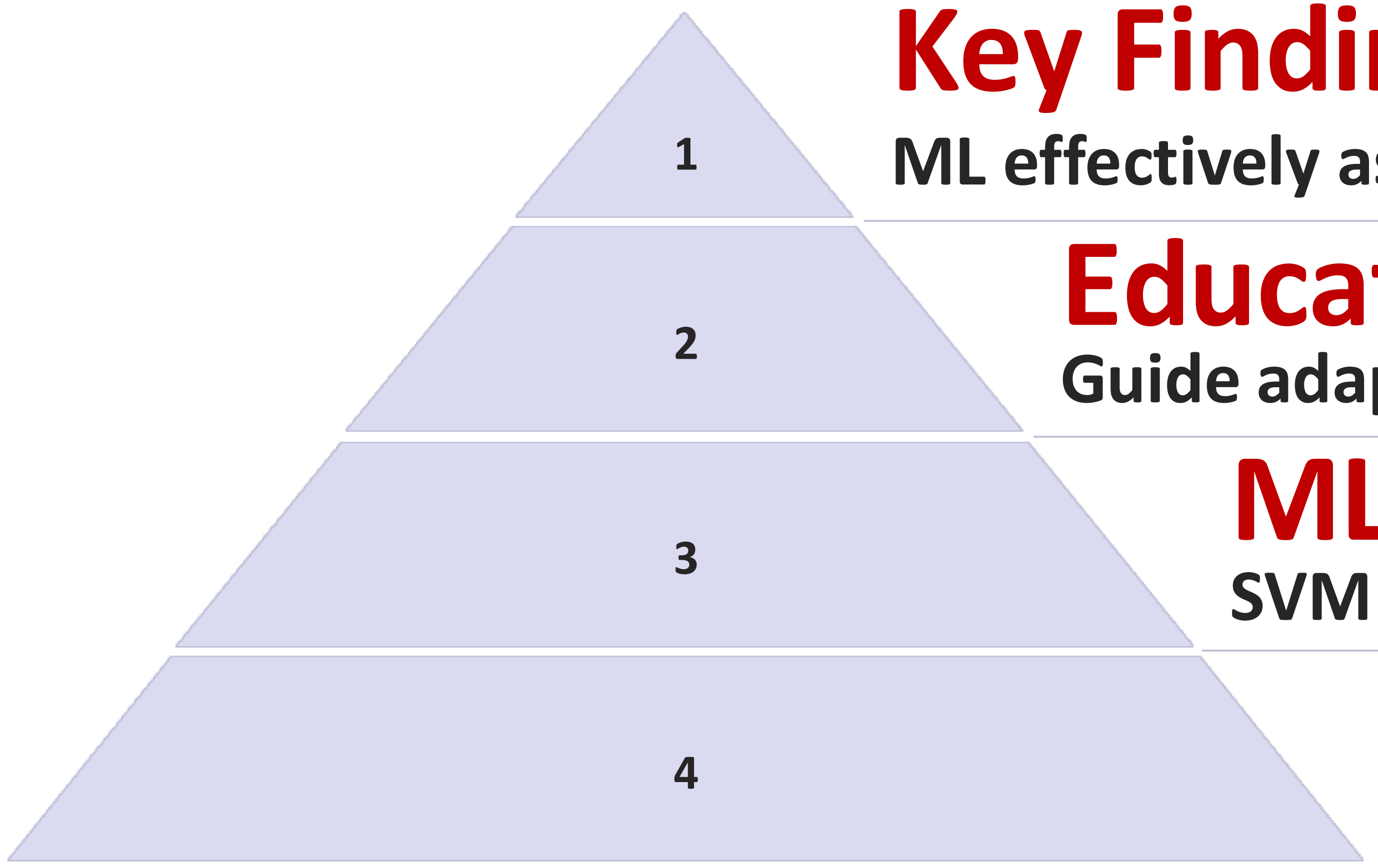
Algorithmic and creative thinking are crucial for CT development

3

Enhance Curriculum

- *Supports data-driven educational policies in engineering training*
- *Findings can guide curriculum adjustments and personalized learning.*

Conclusion



Key Finding
ML effectively assesses CT in engineering

Educational Applications
Guide adaptive learning and interventions

ML Model integration
SVM provides most reliable predictions

Future Directions
Expand dataset and explore deep learning

Questions

Thank you

Contact me:
Opeyemi.ojajuni@sus.edu