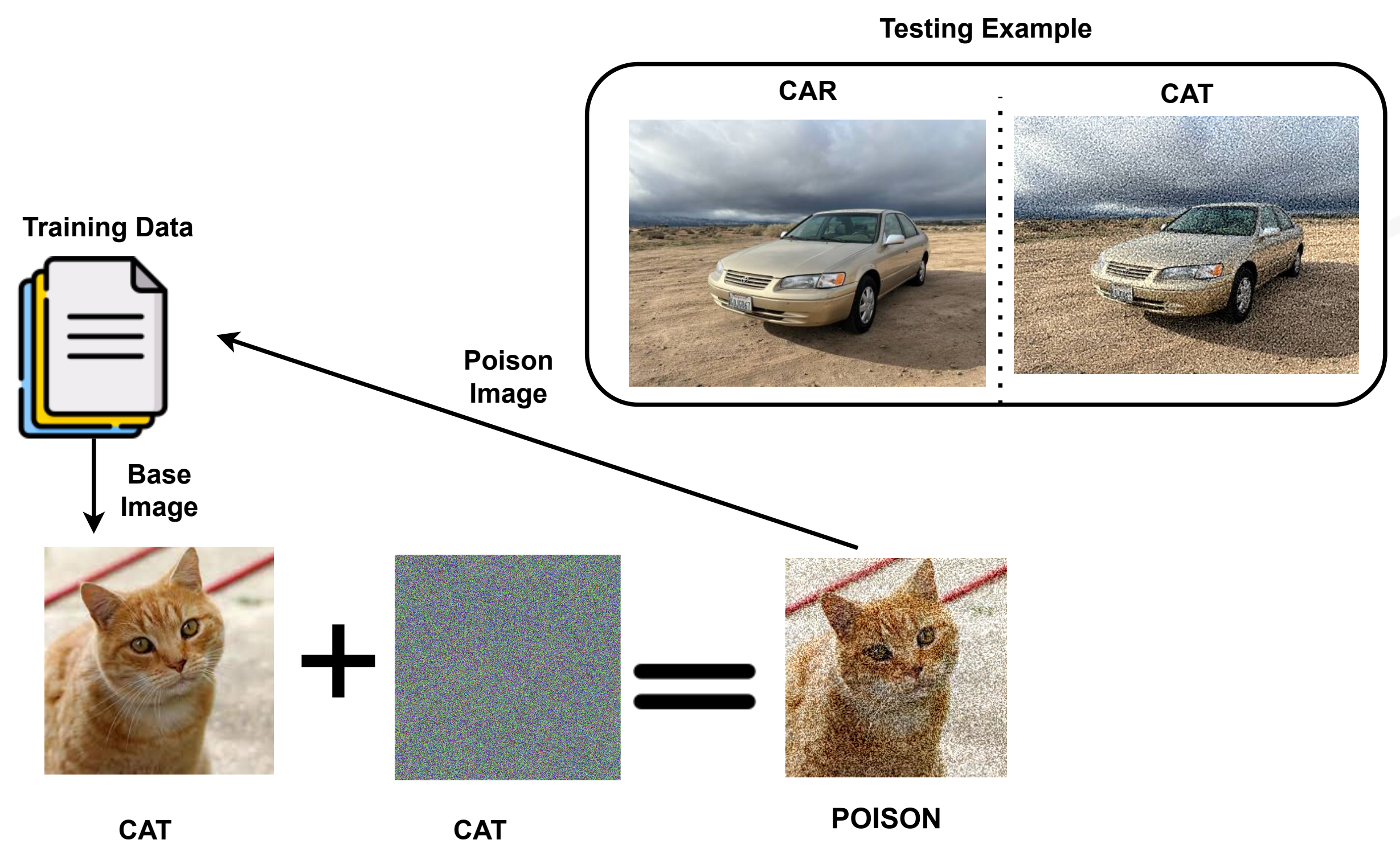
Making Smart Contracts Predict and Scale

Syed Badruddoja, Ram Dantu, Yanyan He, Mark Thompson, Kritagya Upadhyay, Abiola Salau
University of North Texas

Background

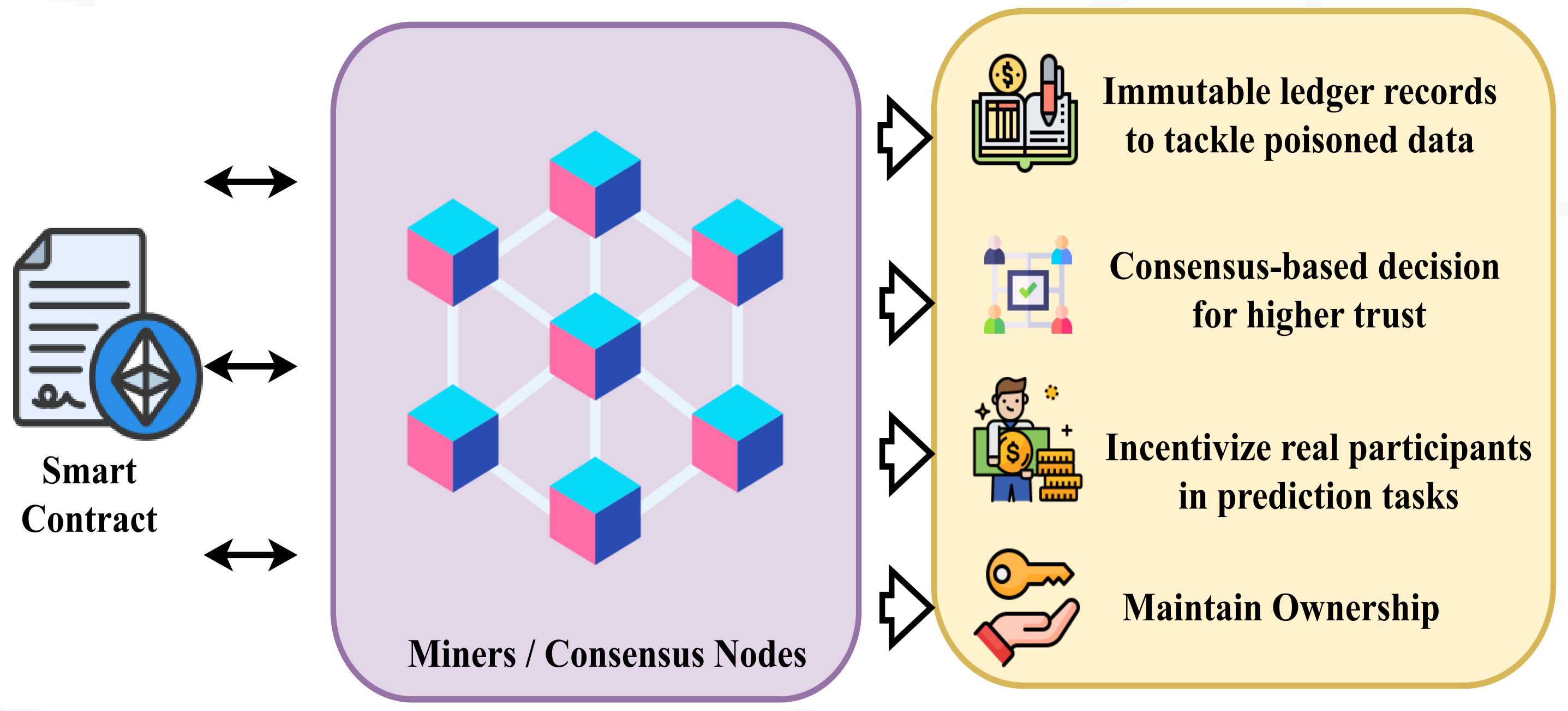
The machine learning algorithms can predict events based on the trained models and datasets. However, a reliable prediction requires the model to be trusted, tamper-resistant and free from poisoning attack. Blockchain technology provides trusted output with consensus-based transactions and an immutable distributed ledger. The machine learning algorithms can be trained on blockchain smart contracts to produce trusted models for reliable prediction. But most smart contracts in the blockchain do not support floating-point data type, limiting computations for classification, which can affect the prediction accuracy.

Model Poisoning Attack



A model can be poisoned by training wrong images as shown above to falsify prediction. As a result, the prediction is falsified and unreliable.

Blockchain Properties



- Immutable ledger records to tackle poisoned data
- Consensus-based decision for higher trust
- Incentivize real participants in prediction tasks
- Maintain Ownership

Blockchain is a distributed ledger of transaction with consensus-based validation which is highly trusted for immutable records in a decentralized network.

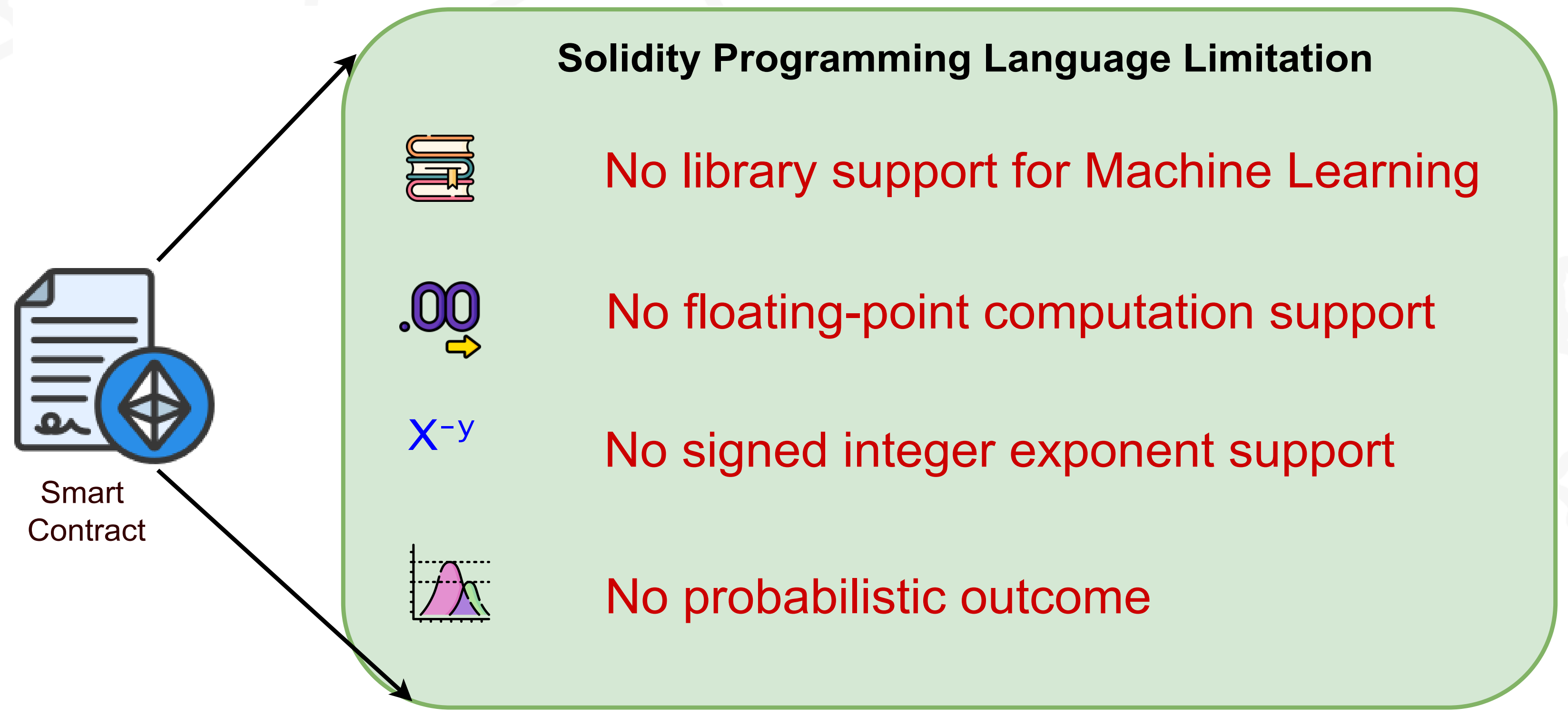
Purpose

- Identify limitations of blockchain to predict classes or events using machine learning algorithm
- Derive numerical methods to overcome limitation and allow prediction in blockchain using machine learning algorithm

Research Questions

- **Can Smart Contracts Compute Probability??**
- **Can Smart Contracts Learn using Machine Learning Algorithms?**
- **Can Smart Contracts Predict the Future?**
- **Can Smart Contracts Predict Accurately?**

Smart Contracts Struggle to Predict



Solidity Programming Language Limitation

- No library support for Machine Learning
- No floating-point computation support
- No signed integer exponent support
- No probabilistic outcome

Smart contracts are used by blockchain network for transactions. Solidity is one of the popular smart contract languages that is highly used by developers. The above limitations prohibits smart contracts to predict using machine learning algorithms.

Solution: Integer-based Probability

We have devised a numerical method based on Taylor's series expansion for converting the fixed-point calculations to integer-based estimation and facilitate smart contracts to classify and predict based on estimated probability. We have chosen naive Bayes algorithm to estimate the probability of a class using only integer operation.

Naive Bayes Formula

Conditional Probability

$$P(c\,|x) = \frac{P(x\,|\,c)\ * P(c)}{P(x)}$$

Expanding conditional probability for all x features

$$P(\,x\,|\,c) * P(c) = (P(x_1\,|\,c)\ * P(x_2\,|\,c) *\ \ldots\ldots\ldots * P(x_n)) * P(c)$$

Gaussian Probability : $P(x\,|c\,) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{\sigma}}$

c is class, x is feature, P is the probability, σ is the variance of features for particular class c, μ is the mean of features for a particular class c.

Comparing Integer Outputs

Derived probability estimation for two classes C_k and C_l

$$\frac{p(C_k)^2}{p(C_l)^2} = \frac{\frac{Ak}{Bk}e^{-\frac{Ck}{Dk}}.(p(C_k))^2}{\frac{Al}{Bl}e^{-\frac{Cl}{Dl}}.(p(C_l))^2} \ldots\ldots..eqn(1)$$

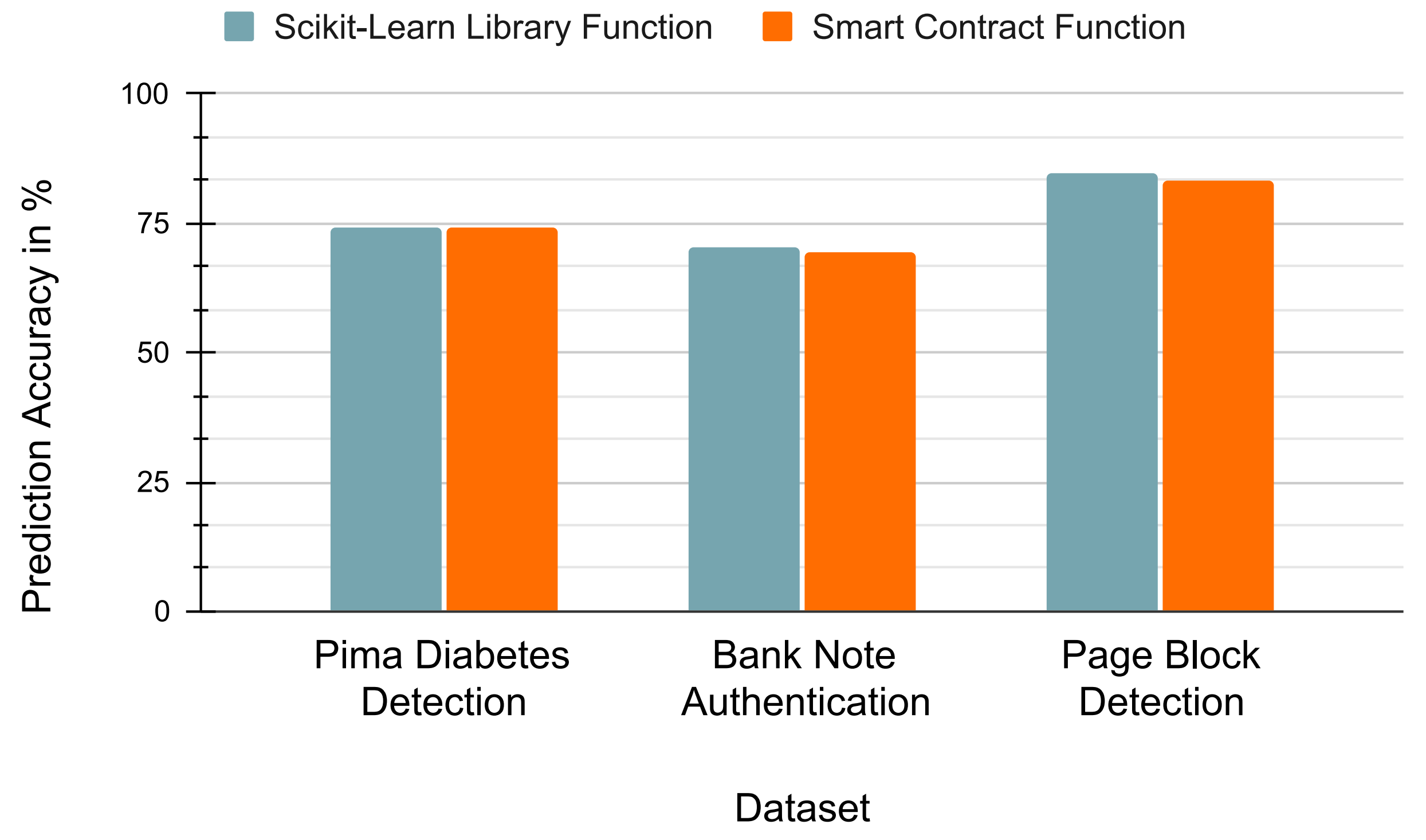
$$\frac{p(C_k)^2}{p(C_l)^2} = \frac{A}{B}.e^{-Q}.e^{-\frac{R}{D}}.\frac{(p(C_k))^2}{(p(C_l))^2} \ldots\ldots\ldots eqn(2)$$

Final two equations of derivations are shown above. From equation (2) we compare probability estimations in integers between individual classes of C_k and C_l to identify the higher probable classes between them. For more classes, similar comparisons are performed until higher probable class is found which will be the predicted class for the input data with unknown class.

Dataset

Dataset	Features	Training Samples	Testing Samples
Pima Diabetes	8	614	154
Bank Note Authentication	4	1097	278
Page Block Detection	10	4378	1095

Prediction Accuracy



Comparable Prediction Accuracy of Smart Contracts against Python Libraries

Reference

S. Badruddoja, R. Dantu, Y. He, M. Thompson, A. Salau and K. Upadhyay, "Making Smart Contracts Predict and Scale," *2022 Fourth International Conference on Blockchain Computing and Applications (BCCA)*, San Antonio, TX, USA, 2022, pp. 127-134, doi: 10.1109/BCCA55292.2022.9922480.

HOTSOS 2023
10TH ANNUAL
HOT TOPICS *in the* SCIENCE OF SECURITY
APRIL 3 - 5, 2023 | *Virtually hosted by* The National Security Agency
hotsos.org