Privacy-Preserving Machine Learning for Collaborative Data Sharing via Auto-encoder Latent Space Embeddings

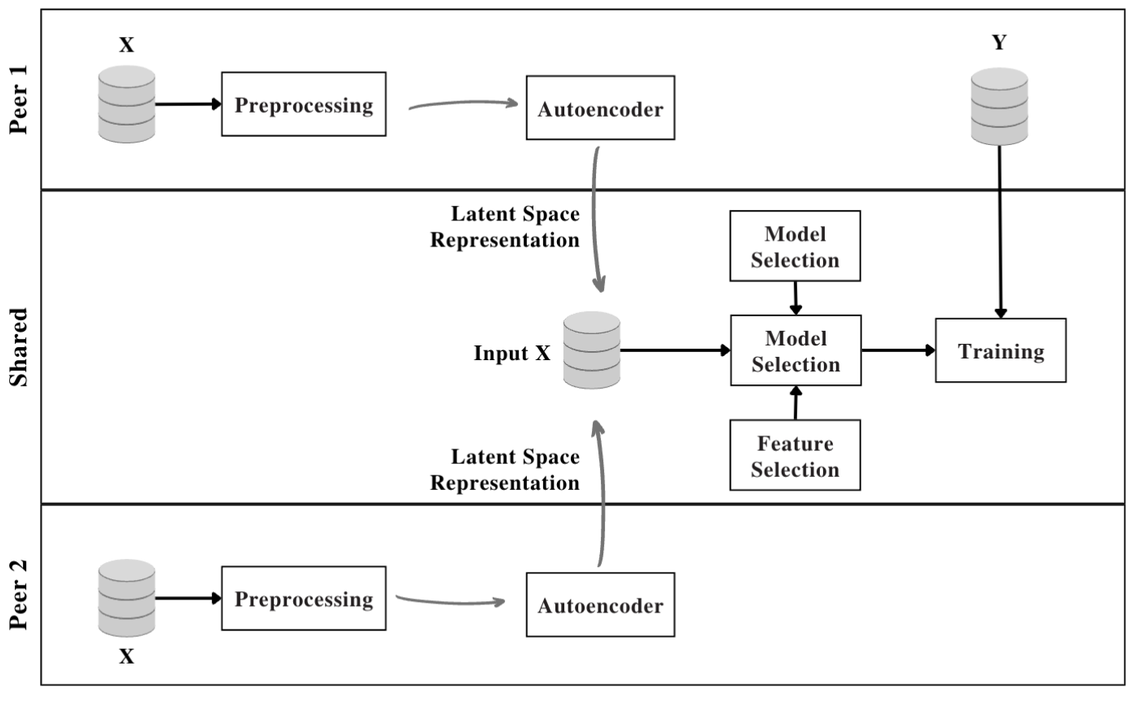
Ana María Quintero-Ossa
Jesús Solano | Hernán Jarcía | David Zarruk | Alejandro Correa | Carlos Valencia

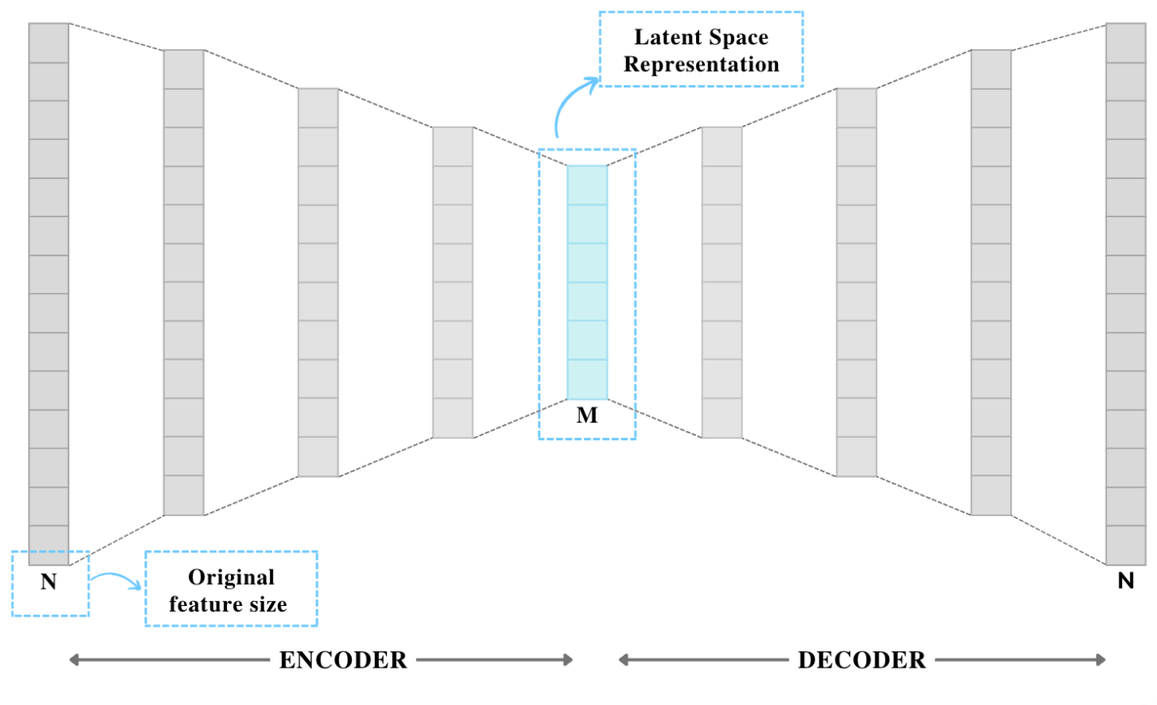
NEURAL INFORMATION PROCESSING SYSTEMS

PROBLEM DEFINITION

This paper presents an innovative framework that uses Representation Learning via autoencoders to generate privacy-preserving embedded data. Thus, organizations can share the data representation to increase machine learning models' performance in scenarios with more than one data source for a shared predictive downstream task.

METHODOLOGY





RESULTS

Scenario 0: Trains a predictive model for the downstream task using a single data source
Scenario 1: we preprocess a unique dataset to obtain a single representation vector and use it to train a predictive model
Scenario 2: Simulate two peers and preprocessing them individually to obtain a representation vector for each source
Scenario 3 and 4: we adjust the autoencoder and transform it into a multitask neural network that, on one side, predicts the representation performance and, on the other, predicts the objective variable

We tested this methodology in 3 case study: House Price, Mnist Numbers, and Buzz in Social Media. These are the results

	Metrics	Scenario 0	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Train	R2	96.17%	91.13%	89.28%	94.21%	89.01%
	MAPE	22.25%	28.46%	32.18%	25.38%	30.27%
Validation	R2	96.14%	91.08%	89.68%	93.87%	88.47%
	MAPE	24.76%	28.51%	33.42%	26.19%	31.06%
Test	R2	96.19%	91.55%	89.01%	94.03%	87.90%
	MAPE	23.87%	28.94%	33.23%	25.94%	31.89%

	Metrics	Scenario 0	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Train	Accuracy	94%	88%	84%	92%	85%
	Precision	94%	88%	84%	92%	85%
	Recall	94%	88%	84%	92%	85%
Validation	Accuracy	92%	88%	84%	91%	84%
	Precision	92%	88%	84%	91%	84%
	Recall	92%	88%	84%	91%	84%
Test	Accuracy	92%	88%	84%	91%	84%
	Precision	92%	88%	84%	91%	84%
	Recall	92%	88%	84%	91%	84%

	Metrics	Scenario 0	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Train	R2	90.26%	84.26%	89.30%	89.79%	88.78%
	MAPE	15.31%	18.03%	16.03%	15.69%	17.74%
Validation	R2	90.32%	84.41%	89.30%	88.93%	87.26%
	MAPE	14.88%	17.91%	15.89%	15.39%	16.89%
Test	R2	90.29%	84.27%	89.33%	89.21%	87.36%
	MAPE	15.09%	17.97%	15.96%	15.27%	17.58%

CONCLUSIONS AND FUTURE WORK

In this paper, we propose an alternate solution to traditional privacy-preserving approaches in machine learning and proof that with an accurate representation learning model, peers can share an embedded dataset that follows the observations' patterns and behavior. Changing the original features to a latent space representation does not drastically deteriorate the performance of the downstream task. In our use cases, the model results decreased for less than 10pp with a representation error between 5% and 11%. Therefore, peers or organizations can collaborate without risking the organization's privacy policies or violating potential clients' privacy concerns.