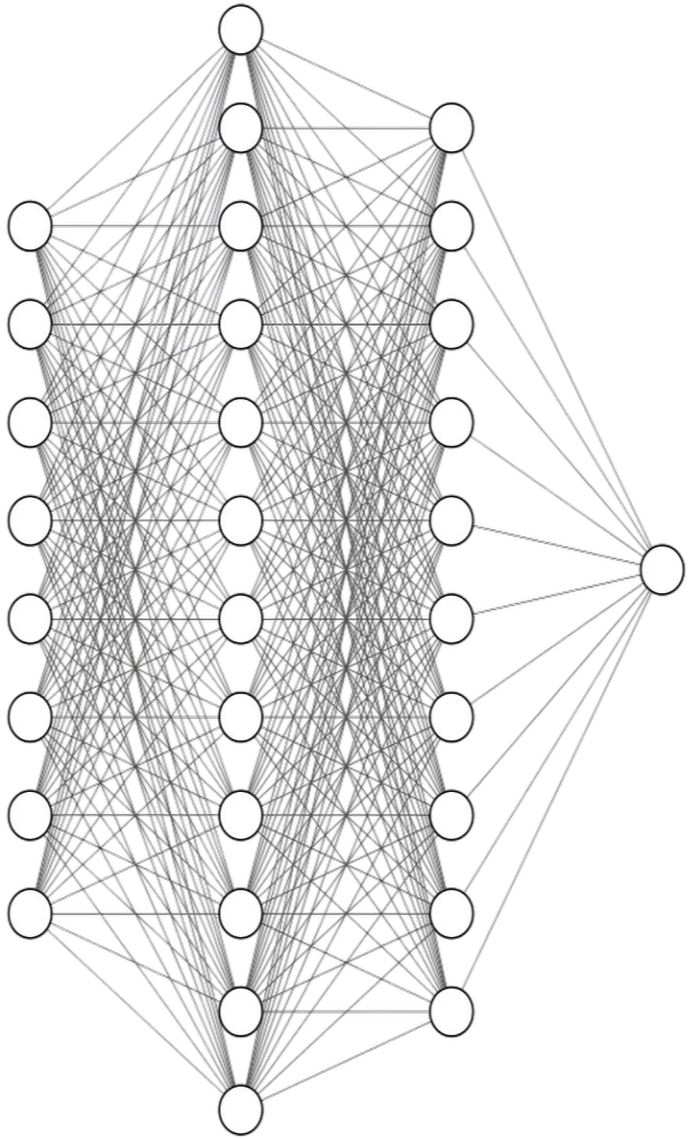
Using Machine Learning for Particle Tracking at the Large Hadron Collider



ENLACE 2023

September 28, 2023

Alejandro Daniel Dennis Hernandez (Tecnológico de Monterrey)

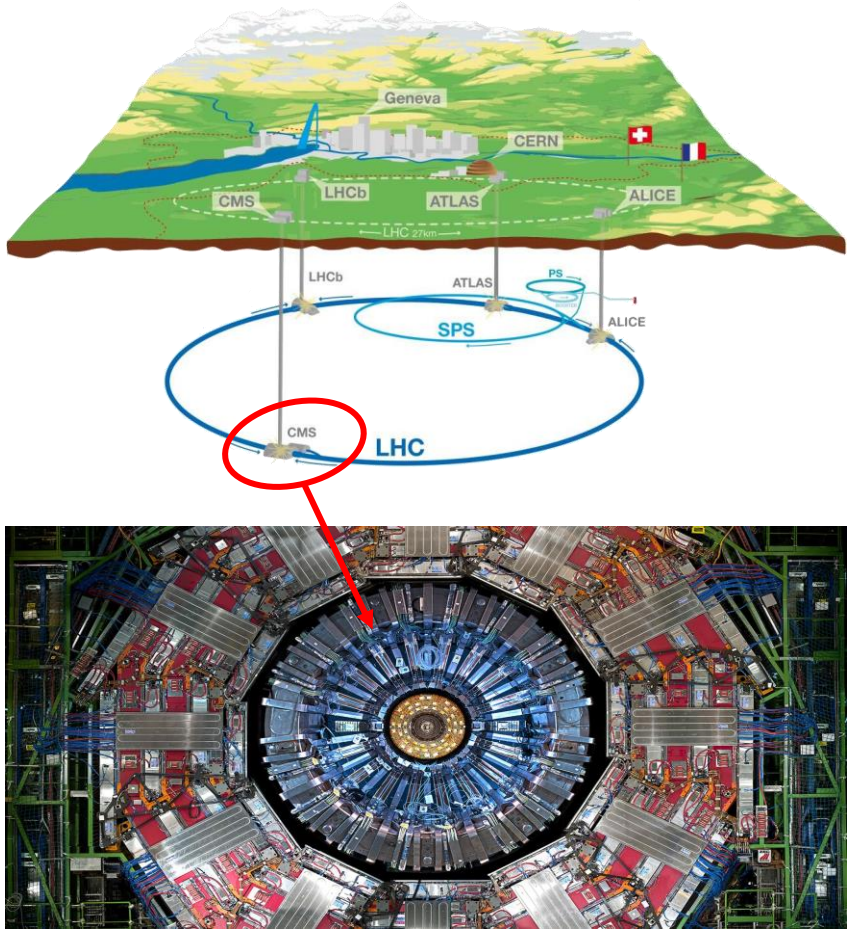
Abraham Jhared Flores Azcona (Instituto Tecnológico de Tijuana)

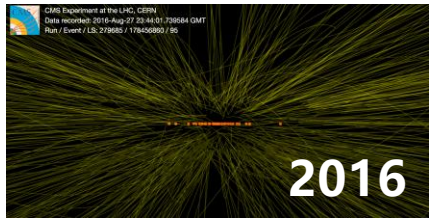
UC San Diego

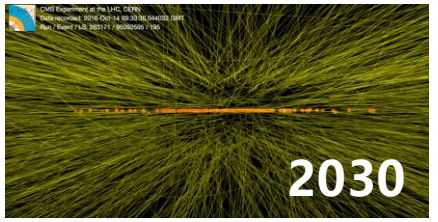
Tecnológico de Monterrey

TIJUANA, BAJA CALIFORNIA

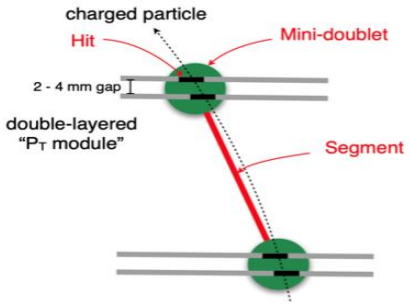
Introduction: LHC & LST



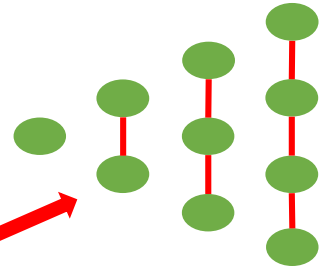




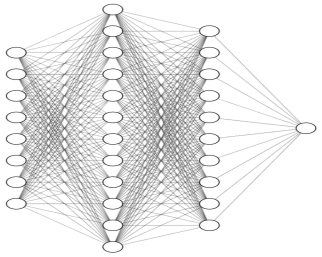
Main Input data: LS



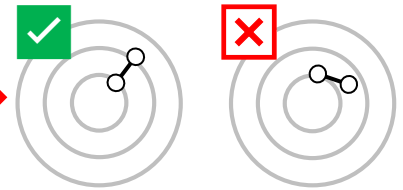
LST: Recreate particle tracks with LS



Support Tool: ML



Objective: distinguish real & fake LS from LST



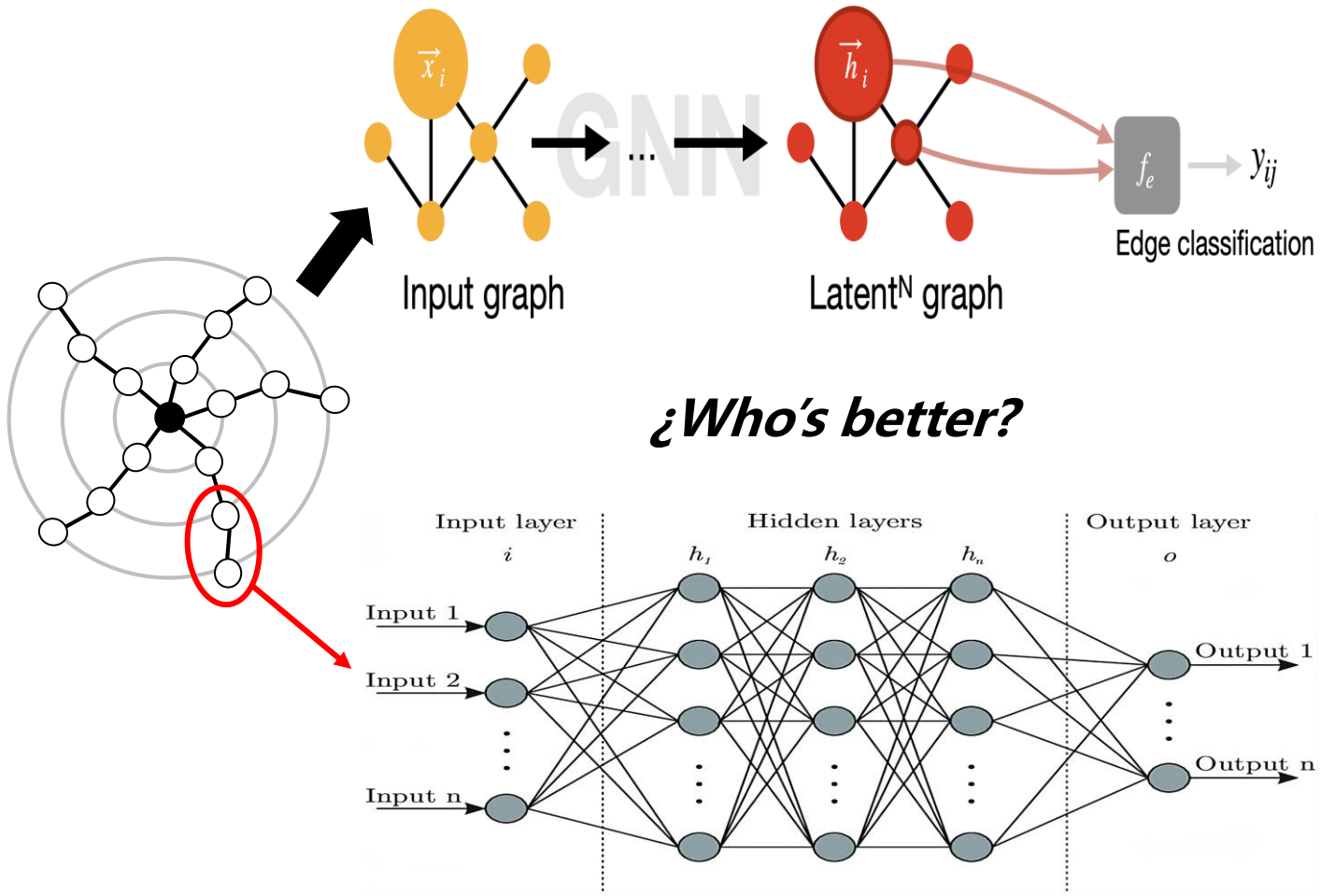
UC San Diego Tecnológico de Monterrey

GNN vs DNN

GNN*: *Graph Neural Network.*
- **7 node features & 3 edge features**
- **1 hidden layer**
- **200 neurons per layer**
- **Learning Rate of 0.005, *stepped down by a factor of 0.7 every 5 epochs***
- **50 epochs**

Big DNN:** *Big Deep Neural Network.*
- **17 input features**
- **2 hidden layers**
- **200 neurons per layer**
- **Learning Rate of 0.002**
- **100 epochs**

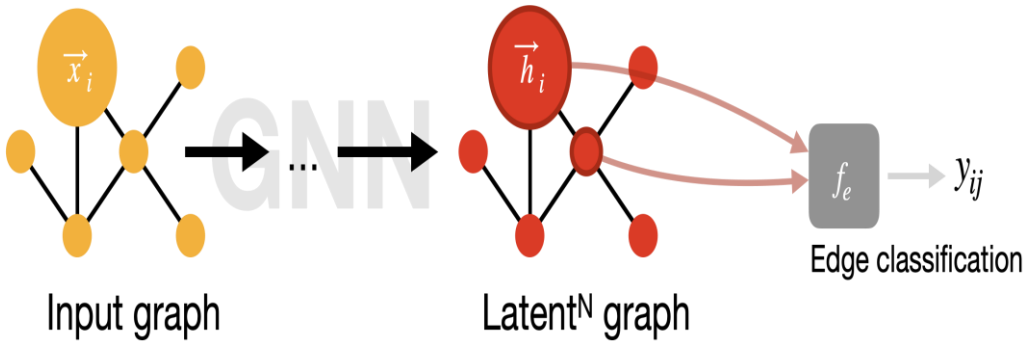
Small DNN: *Small Deep Neural Network.*
- **17 input features**
- **2 hidden layers**
- **32 neurons per layer**
- **Learning Rate 0.002**
- **50 epochs**



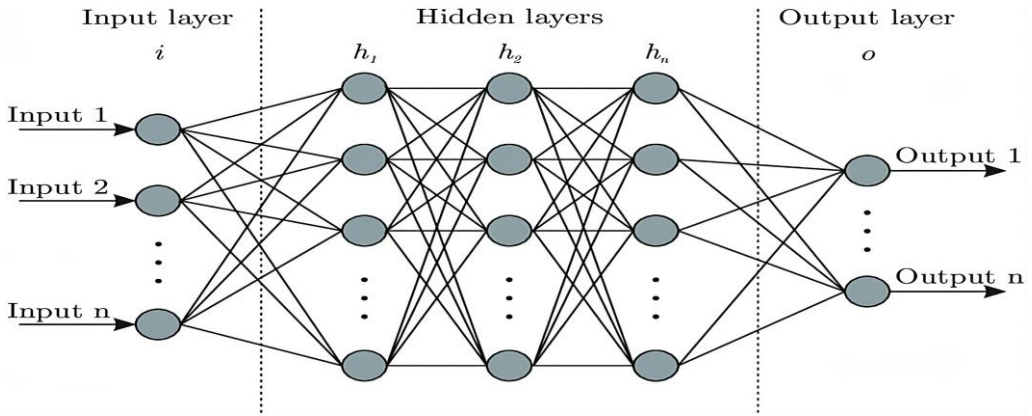
***Trained by Phillip Chang**
****The rest of the slides focus on this model**

GNN vs DNN

We considered *a-priori* that **the performance of the GNN will be greater than the Big DNN and Small DNN**



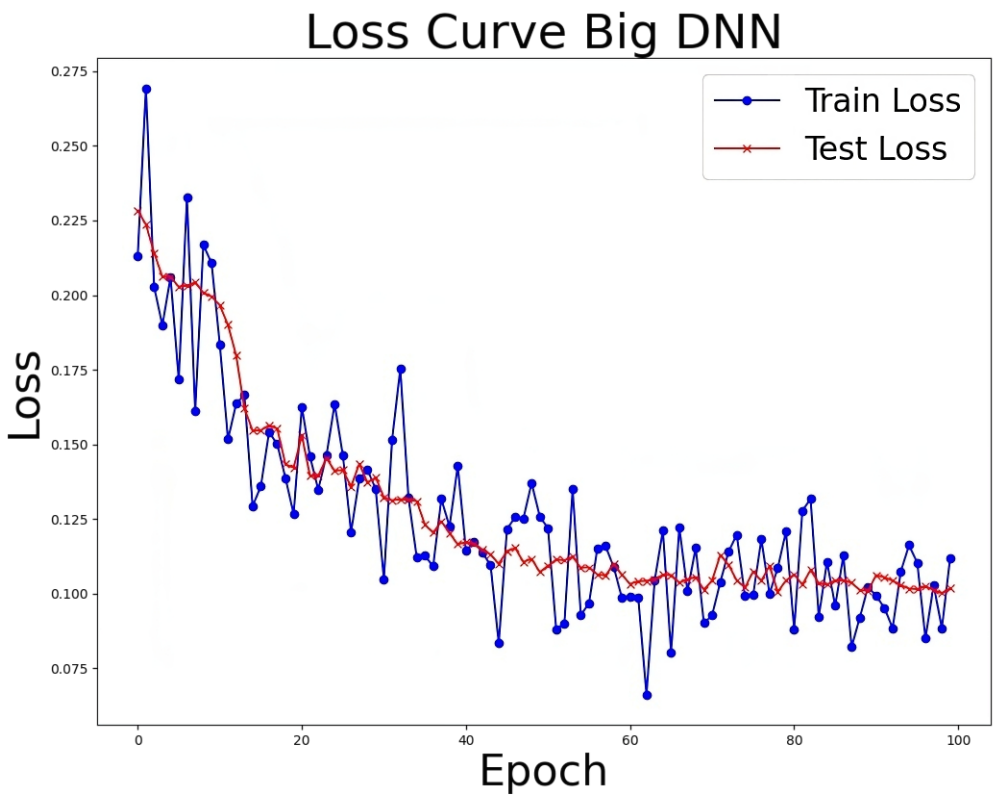
GNN >> Big DNN
GNN >> Small DNN

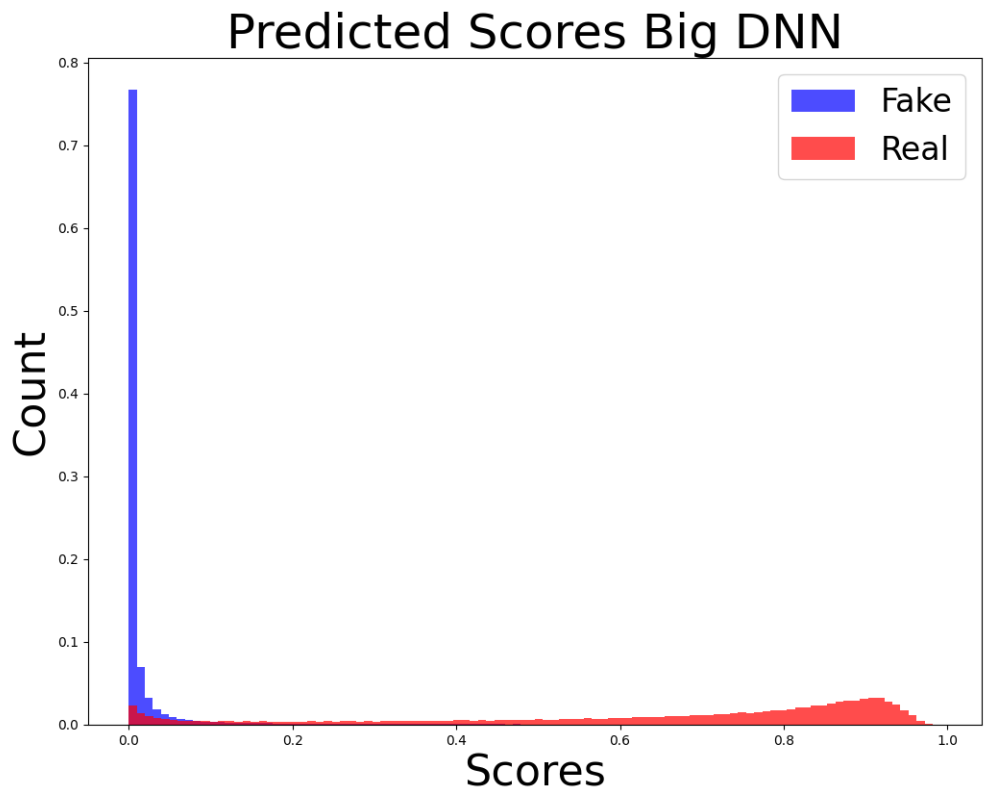


UC San Diego Tecnológico de Monterrey

Training of the Big DNN

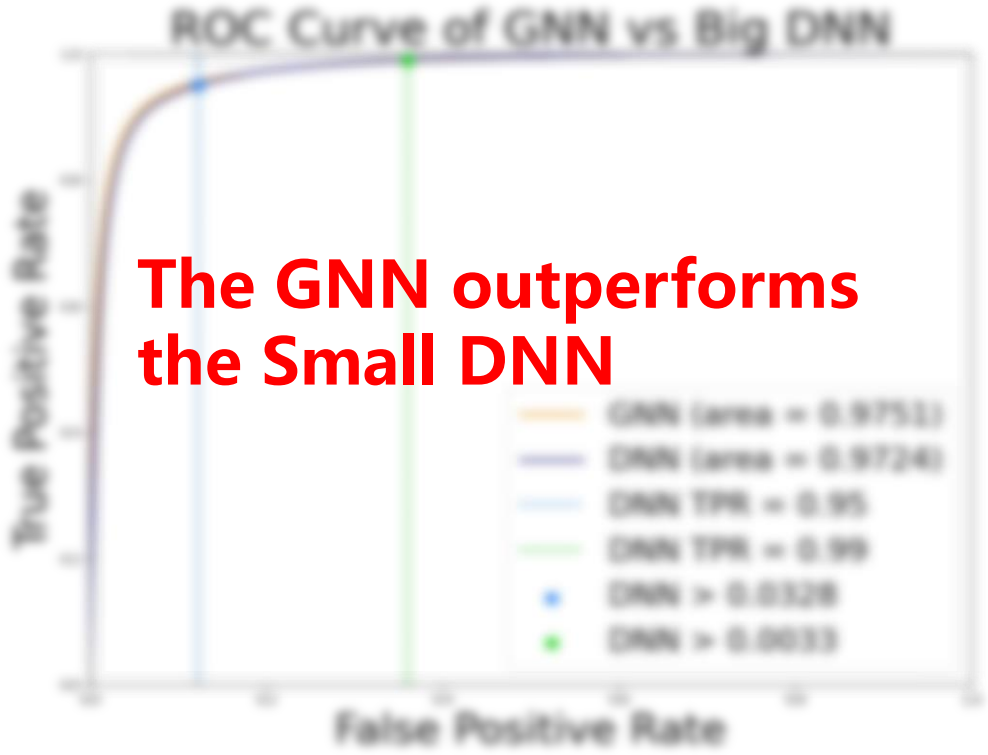
Key Takeaway: **The DNN does not overfit and it's effectively distinguishing fake from real LS.**



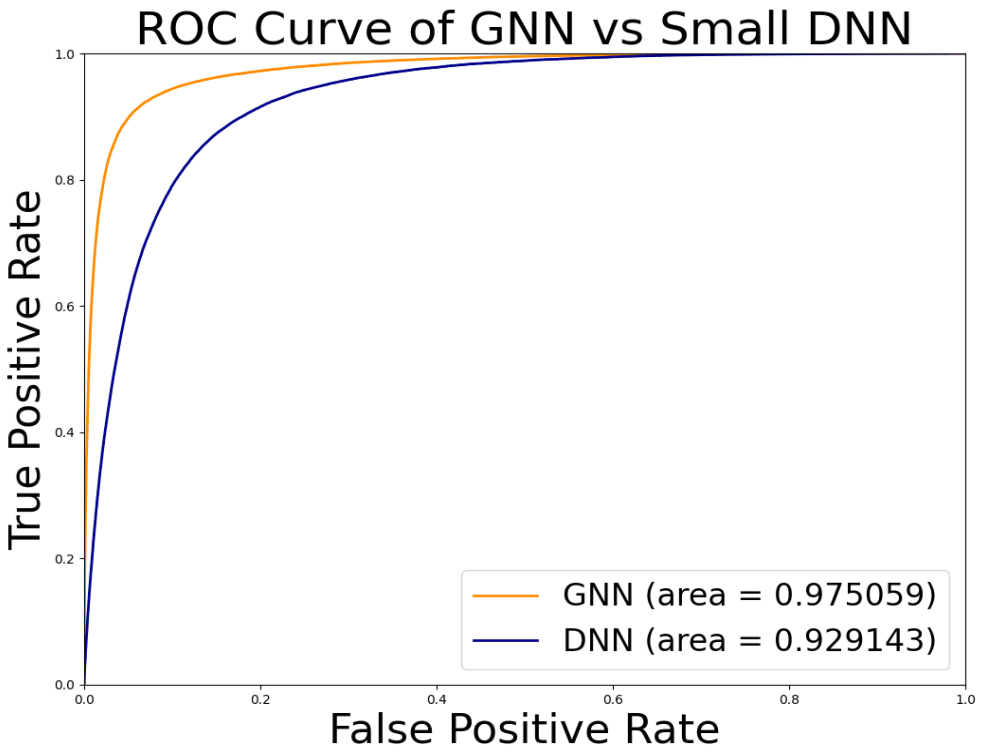


UC San Diego Tecnológico de Monterrey

Results

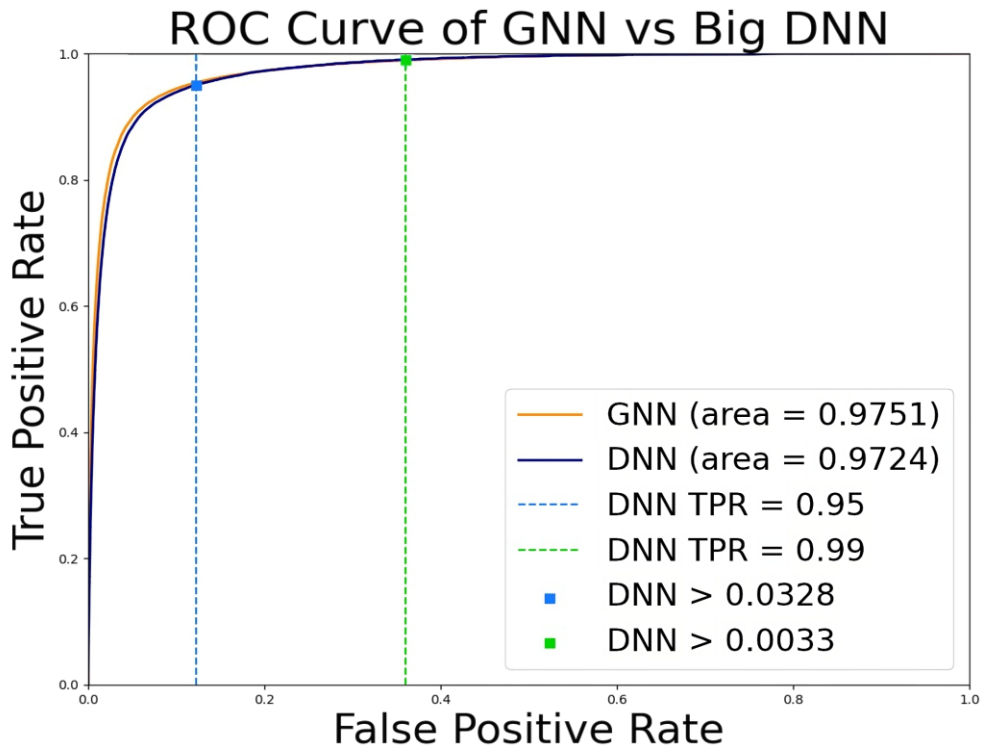


The GNN outperforms the Small DNN



UC San Diego Tecnológico de Monterrey

Results

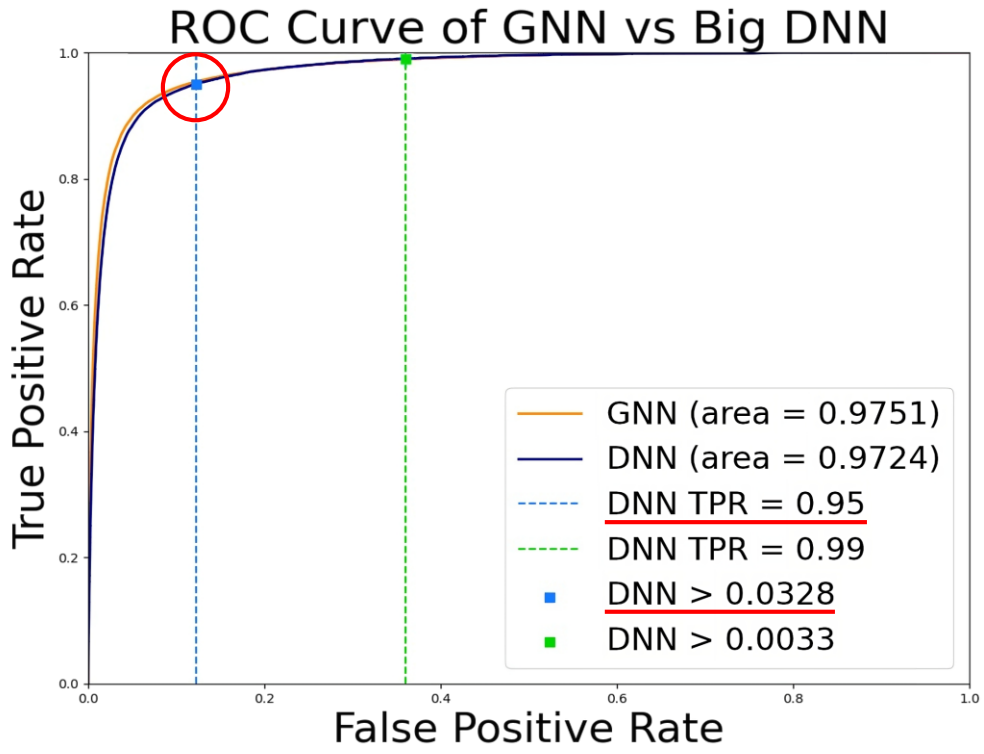


The Big DNN achieves similar performance to the GNN although being a much simpler architechture.

UC San Diego

Tecnológico de Monterrey

Results



LS Predictions when TPR = 0.95

	GNN > X	**DNN > Y**	**GNN ∩ DNN**
Real	49628	49628	48966
Fake	78847	85248	57704
Σ	128475	134876	106670

Nota: **X**, **Y** are threshold values that satisfy TPR = 0.95 where **X** = 0.0328, **Y** = 0.0385. **Σ** = **Real** + **Fake.** TPR: True Positive Rate. LS: Line Segments.

UC San Diego Tecnológico de Monterrey

Results

Real LS with unique ID when TPR = 0.95

Note: The rings enclose the addition of the **GNN ∩ DNN (48966)**, and **GNN** and the **DNN.** The previous result lets us calculate **Real GNN (49628)** y **Real DNN (49628)**.

LS Predictions when TPR = 0.95

	GNN > X	DNN > Y	GNN ∩ DNN
Real	49628	49628	48966
Fake	78847	85248	57704
Σ	128475	134876	106670

Note: **X**, **Y** are threshold values that satisfy TPR = 0.95 where **X** = 0.0328, **Y** = 0.0385. **Σ** = **Real** + **Fake.** TPR: True Positive Rate. LS: Line Segments.

Results

Fake LS with unique ID when TPR = 0.95

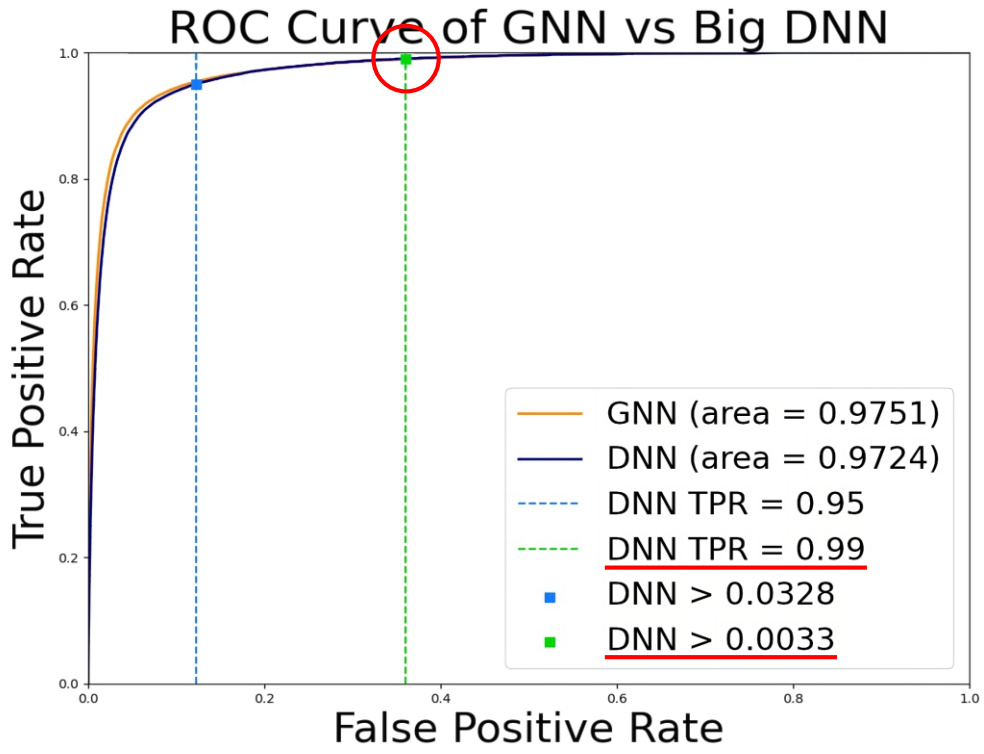
Note: The rings enclose the addition of the **GNN ∩ DNN (57704)**, and **GNN** and the **DNN.** The previous result lets us calculate **Fake GNN (78847)** y **Fake DNN (85248)**.

LS Predictions when TPR = 0.95

	GNN > X	DNN > Y	GNN ∩ DNN
Real	49628	49628	48966
Fake	78847	85248	57704
Σ	128475	134876	106670

Note: **X**, **Y** are threshold values that satisfy TPR = 0.95 where **X** = 0.0328, **Y** = 0.0385. **Σ** = **Real** + **Fake.** TPR: True Positive Rate. LS: Line Segments.

Results



LS Predictions when TPR = 0.99

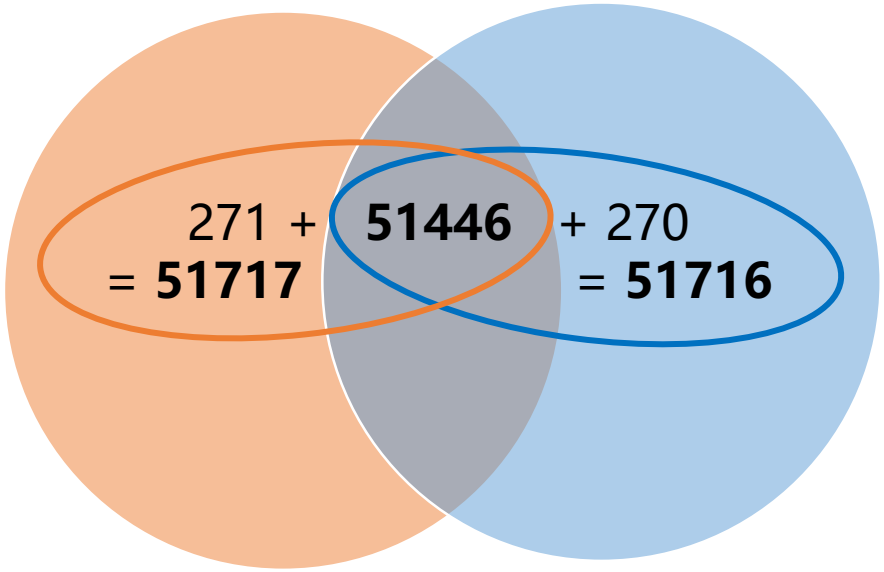
	GNN > X	**DNN > Y**	**GNN ∩ DNN**
Real	51717	51716	51446
Fake	261839	250781	193761
Σ	313556	302497	245207

Nota: **X**, **Y** are threshold values that satisfy TPR = 0.99 where **X** = 0.0033, **Y** = 0.0045. **Σ** = **Real** + **Fake.** TPR: True Positive Rate. LS: Line Segments.

UC San Diego Tecnológico de Monterrey

Results

Real LS with unique ID when TPR = 0.99



Note: The rings enclose the addition of the **GNN ∩ DNN (51446)**, and the **GNN** and the **DNN.** The previous result lets us calculate **Real GNN (51717)** y **Real DNN (51716)**.

LS Predictions when TPR = 0.99

	GNN > X	DNN > Y	GNN ∩ DNN
Real	51717	51716	51446
Fake	261839	250781	193761
Σ	313556	302497	245207

Note: **X**, **Y** are threshold values that satisfy TPR = 0.99 where **X** = 0.0033, **Y** = 0.0045. **Σ** = **Real** + **Fake.** TPR: True Positive Rate. LS: Line Segments.

Results

Fake LS with unique ID when TPR = 0.99

Note: The rings enclose the addition of the **GNN ∩ DNN (193761)**, and the **GNN** and the **DNN.** The previous result lets us calculate **Fake GNN (261839)** y **Fake DNN (250781)**.

LS Predictions when TPR = 0.99

	GNN > X	DNN > Y	GNN ∩ DNN
Real	51717	51716	51446
Fake	261839	250781	193761
Σ	313556	302497	245207

Note: **X**, **Y** are threshold values that satisfy TPR = 0.99 where **X** = 0.0033, **Y** = 0.0045. **Σ** = **Real** + **Fake.** TPR: True Positive Rate. LS: Line Segments.

Conclusion

- **The Big DNN achieves similar performance to the GNN.**
- Considering the technical debt of the compared models, **the Big DNN is the best option to classify LS.**
- **Further research is needed to validate our results.**

Acknowledgements

- Dra. Olivia. A. Graeve (Directora ENLACE 2023)
- Dr. Hector Arias (Coordinador ENLACE-ITT)
- José Guillermo Cárdenas López (Director ITT)
- Alexia Chavez, Diego Velazquez y Brandon A. Meza (Organizadores ENLACE-ITT y compañeros de ENLACE 2023)
- Solar Turbines Tijuana (Patrocinador de Beca ENLACE 2023)
- Dr. Frank Wuertwein (PI)
- Jonathan Guiang (Mentor)

UC San Diego Tecnológico de Monterrey