

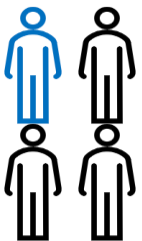
# The Cost of Stroke in Canada



About  
**50,000**  
people suffer strokes in Canada  
each year



After age  
**55+**  
risk of stroke increases rapidly



Approximately  
**1 in 4**  
living with stroke are < age 65

Attributable cost of stroke about

**\$30,000/yr** ↗

per person in Ontario, Canada

*Identifying stroke risk factors is  
essential for improving patient  
outcomes and reducing costs  
on health care system*

# ML-9 Team Project: Stroke Prediction

## Dataset

11 features, 5110 records:

- ✓ Demographic
- ✓ Residence type
- ✓ Health indicators
- ✓ Socioeconomic
- ✓ Stroke occurrence

## Aims



Baseline  
models



Neural  
networks



Evaluation  
metrics

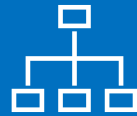


Identify key  
indicators

## Techniques & Technologies



Logistic  
Regression



Random  
Forest



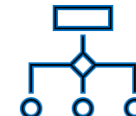
Neural  
Network  
(FCNN)

- Parameter tuning
- Regularization
- Cross-validation

## Risks & Unknowns

- Class imbalance (limited stroke data)
- Missing data, general quality
- Generalizability
- Model complexity

## Outcomes



Stroke-risk  
model



Targeted  
screening



Actionable  
insight



Prioritize  
care

# ML-9 Team Project: Methods

---

- Examine Dataset for relationships, both linear (Logistic Regression) and non-linear (Random Forest)
- Train Fully Connected Neural Network using Keras and/or PyTorch
- Make use of GitHub Organizations Projects tool to drive production
- Use Google Colab to run ML models more efficiently
- Balance data using SMOTE



# ML-9 Team Project: Results

## Logistic Regression Performance Summary

Metric	Class 0 (No Stroke)	Class 1 (Stroke)	Overall
Precision	0.87	0.85	
Recall	0.86	0.86	
F1-score	0.86	0.86	
Accuracy	—	—	0.86
ROC-AUC	—	—	0.94

## Random Forest Performance Summary

Metric	Class 0 (No Stroke)	Class 1 (Stroke)	Overall
Precision	0.96	0.96	
Recall	0.96	0.96	
F1-score	0.96	0.96	
Accuracy	—	—	0.96
ROC-AUC	—	—	0.99

## FCNN Performance Summary

Metric	Class 0 (No Stroke)	Class 1 (Stroke)	Overall
Precision	0.90	0.90	
Recall	0.91	0.89	
F1-score	0.90	0.90	
Accuracy	—	—	0.90
ROC-AUC	—	—	0.97

## Model Comparison

Model	Accuracy	ROC-AUC	F1-score	Interpretability	Complexity
Logistic Regression	0.86	0.94	0.86	✔ High	● Low
Random Forest	0.96	0.99	0.96	⚠ Moderate	● Medium
FCNN	0.90	0.97	0.90	✖ Low	● High

# ML-9 Team Project: Conclusions



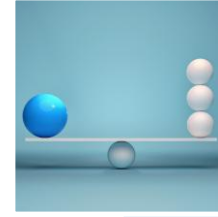
## Performance

- Random Forest achieved best raw metrics



## Interpretability

- Random Forest (with SHAP) offer actionable insights



## Complexity vs. Gain

- Random Forest is most practical