

**RCIF Applied HPC
Seminar Series**

**Model
Evaluation with
TensorBoard for
Healthcare
Applications**



Washington
University in St. Louis

SCHOOL OF MEDICINE

MIR Mallinckrodt Institute
of Radiology

Introduction to Model Evaluation

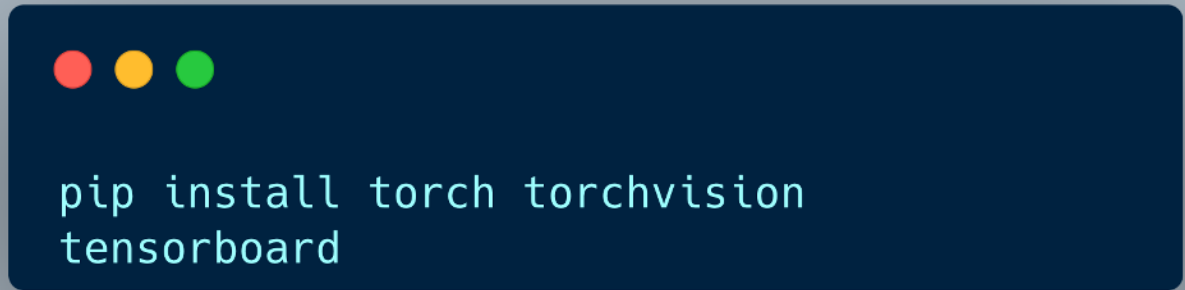
- Model evaluation is crucial in healthcare AI to ensure reliability and safety
- Evaluation helps assess model performance, generalizability, and potential biases
- We'll focus on TensorBoard for visualization
- PyTorch will be our main framework for model development

What's Tensorboard?

- Open-source tool for visualizing machine learning models.
- Provides real-time feedback on training metrics, model architecture, and hyperparameters.
- Helps in debugging and improving model performance during training.
- Supports various visualizations:
 - Scalars (e.g., loss, accuracy)
 - Graphs (model structure)
 - Histograms (weights, biases)
 - Images (input data and output comparisons)
 - Embeddings (dimensionality reduction for high-dimensional data)
- Enables collaboration and transparency with shareable logs and dashboards.



Setting Up Your Environment



```
pip install torch torchvision  
tensorboard
```

A terminal window with a dark blue background and three colored window control buttons (red, yellow, green) in the top left corner. The text is displayed in a light blue monospaced font.

Key Metrics for Healthcare Models


- Accuracy: Overall correctness of predictions
- Precision: Proportion of true positives among positive predictions
- Recall: Proportion of true positives among all actual positives
- F1-score: Harmonic mean of precision and recall
- AUC-ROC: Model's ability to distinguish between classes
- Confusion Matrix: Table showing correct and incorrect classifications

Analyzing Model Predictions

- Evaluate Model Accuracy: Track metrics like accuracy, precision, recall, and F1-score during training and validation.
- Compare Predictions: Visualize predicted vs. actual outcomes for different patient or data segments.
- Confusion Matrix: Use TensorBoard to plot confusion matrices for classification tasks to identify areas of misclassification.
- Precision-Recall Curves & ROC Curves: Evaluate model performance with detailed insights into true positive and false positive rates.
- Analyze Bias: Track performance across different demographic or medical groups to identify biases in model predictions.
- Sample Prediction Visualizations: Display individual predictions with corresponding labels, such as medical image classifications or segmented tumor regions.
- Monitor Custom Metrics: Set up healthcare-specific evaluation metrics (e.g., survival probability, disease progression) for detailed prediction analysis.
- Time-Series Prediction Tracking: Visualize changes in predictions over time for longitudinal healthcare data.



Logging Metrics with PyTorch and TensorBoard



```
from torch.utils.tensorboard import SummaryWriter

writer = SummaryWriter('runs/experiment_1')
for epoch in range(num_epochs):
    writer.add_scalar('Loss/train', train_loss, epoch)
    writer.add_scalar('Accuracy/train', train_accuracy, epoch)
writer.close()
```

Visualizing Model Architecture

- Helps understand model complexity
- Identify potential bottlenecks or errors in architecture
- Useful for communicating model design to clinical collaborators

Hyperparameter Tuning

- TensorBoard HParams dashboard:
- Supports various search algorithms (random, grid, Bayesian)







```
writer.add_hparams({'lr': 0.1, 'bsize': 32}, {'accuracy': 0.95})
```

Model Performance Visualization

- Learning curves: Plot train vs. validation metrics
- Confusion matrix: Use `sklearn.metrics.plot_confusion_matrix`
- ROC curves: `sklearn.metrics.plot_roc_curve`
- Visualize in TensorBoard or export for custom plots



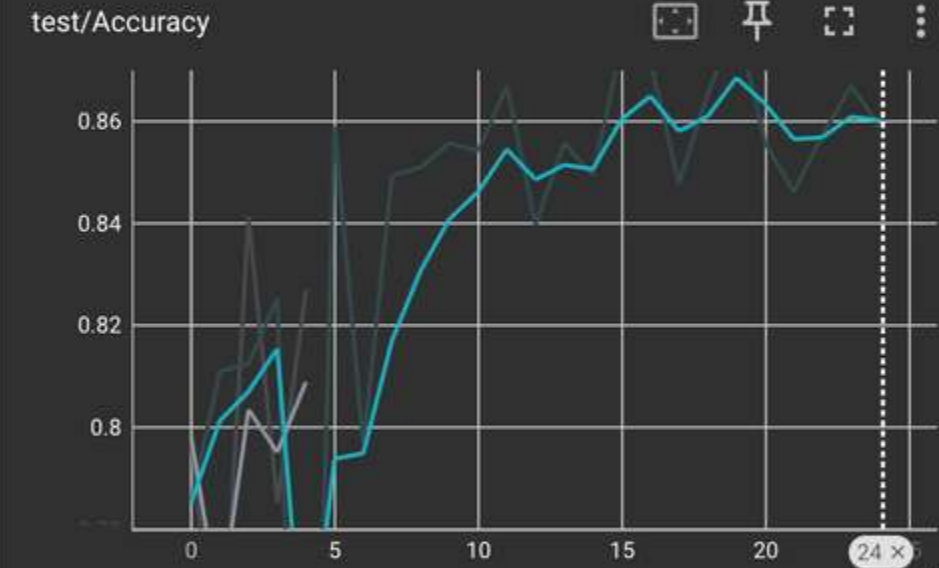
Filter runs (regex)


- ☒ Run ↑ 
- ☒ Sep18_02-08-25_9ba0ce9c652b 
- ☒ Sep18_02-14-35_9ba0ce9c652b 
- ☒ Sep18_03-06-43_9ba0ce9c652b 

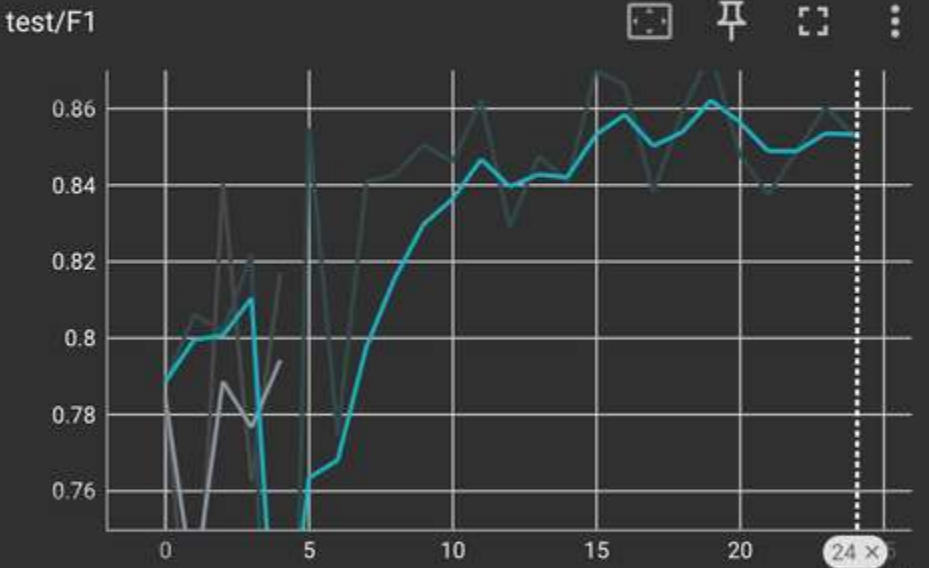
Filter tags (regex)

test 5 cards

All Scalars Image Histogram  Settings




Run ↑	Smoothed	Value	Step	Relative
 Sep18_02-08-25_9ba0ce9c652b	0.8089	0.8269	4	4.538 min
 Sep18_02-14-35_9ba0ce9c652b	0.8601	0.859	24	50.07 min



Run ↑	Smoothed	Value	Step	Relative
 Sep18_02-08-25_9ba0ce9c652b	0.7942	0.8171	4	4.538 min
 Sep18_02-14-35_9ba0ce9c652b	0.8533	0.8529	24	50.07 min

Card Width

☒ Enable saving pins (Scalars only) 

SCALARS

Smoothing

 0.6

Tooltip sorting method

Alphabetical

☒ Ignore outliers in chart scaling

☐ Partition non-monotonic X axis 

HISTOGRAMS

Mode

**Request an
Account!**



**Need More
Help?**



**Sign Up for
the next
Seminar**

