Uncertainty Quantification in Deep Neural Network Models for the Determination of Gene Mutation Status in Gliomas

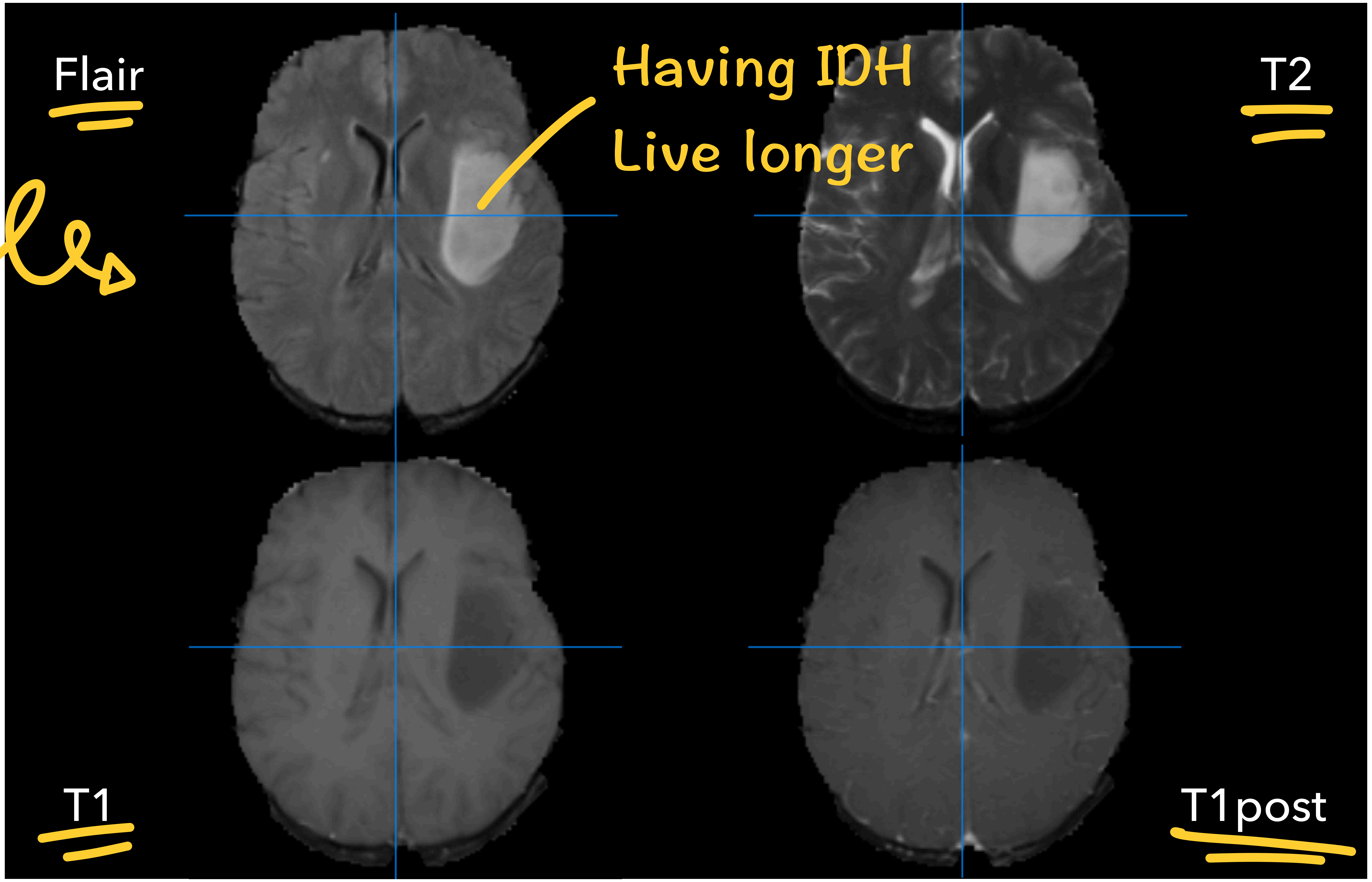
Dingkun Guo

Instructor: Xun Huan

MEUS WI 2019

INTRODUCTION — Medical Background

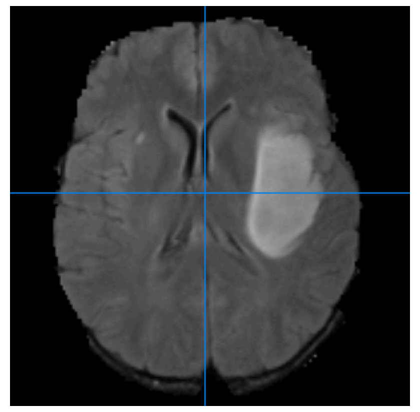
- **Magnetic resonance imaging (MRI)** scans of the patient brain
  - Four types of images are generated with different settings



- **Isocitrate dehydrogenase (IDH) mutation status**
  - Increases survival probability
  - Critical for treatment planning

INTRODUCTION — Model

Input: MRI scans of the brain



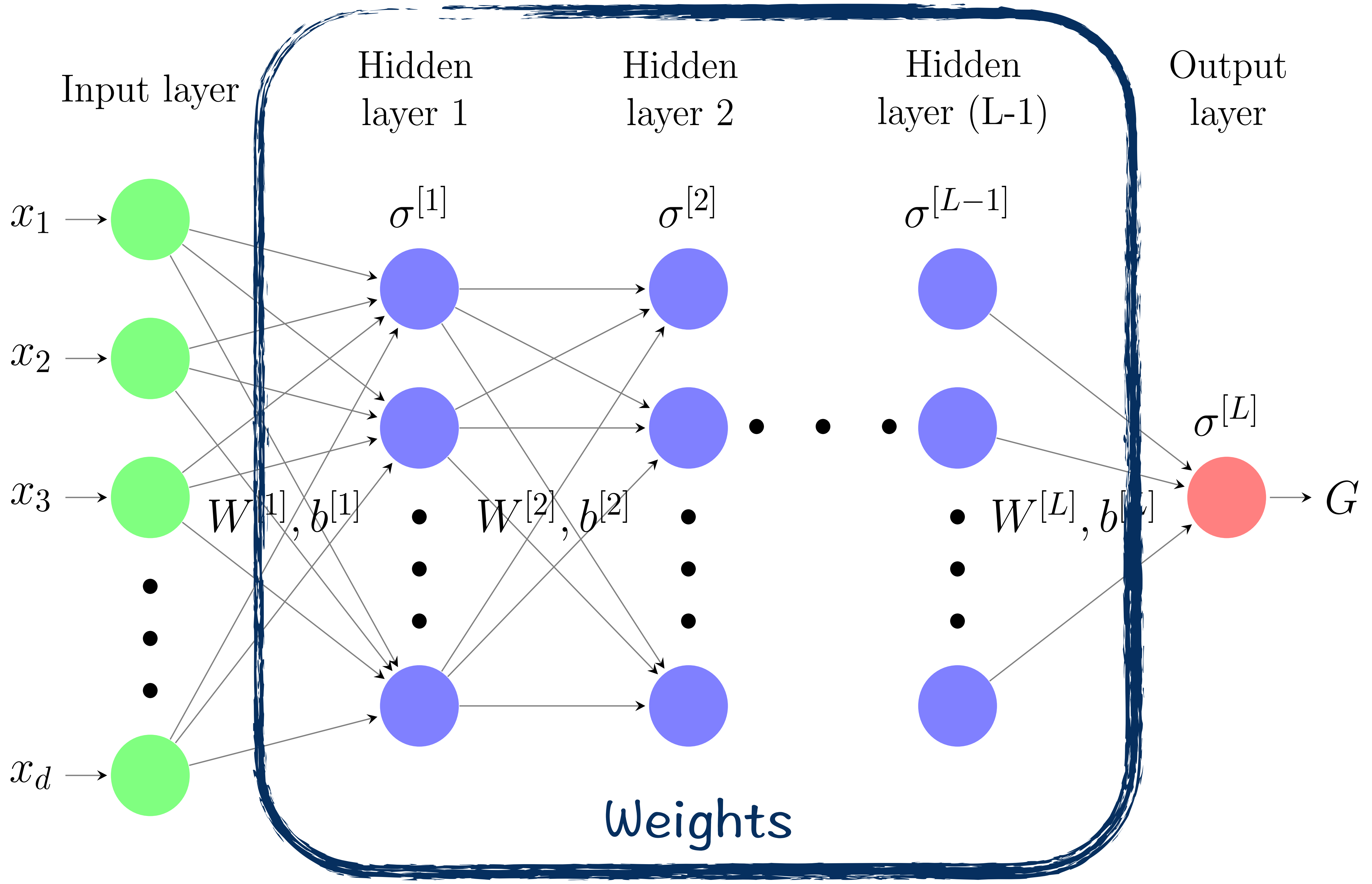
Deep neural networks were developed to predict IDH status from MRI scans [Chang et al., 2018]

Model

80%

Output: prediction probability of positive IDH mutation

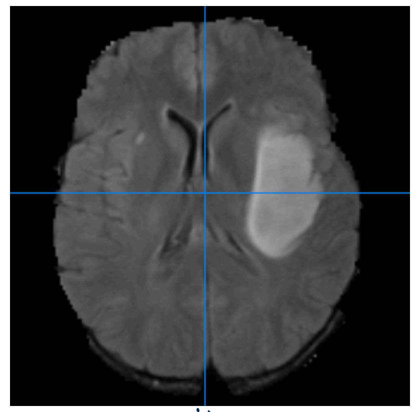
INTRODUCTION — Engineering Background



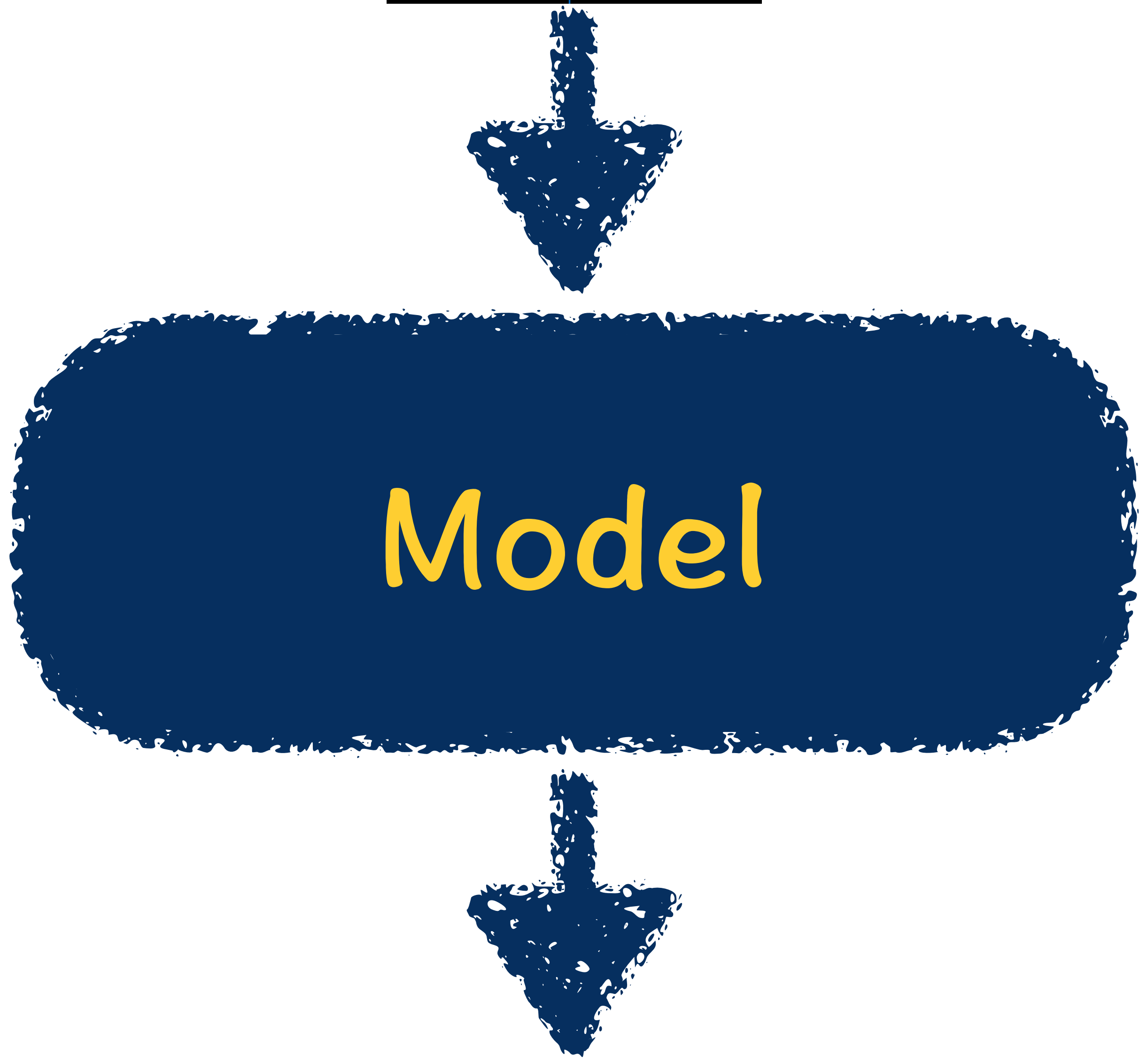
Material
Manufacturing
Atmosphere

- **Deep neural network (DNN)**, a common *data-driven model*, is trained to find relationship between input and output
- **Uncertainty quantification (UQ)** uses statistical methods to produce a measure of confidence on these predictions

INTRODUCTION — Model



Input: MRI scans of the brain



UNCERTAINTY QUANTIFICATION IS CRUCIAL

Deep neural networks were developed to predict IDH status from MRI scans [Chang et al. 2018]

- Over 22 million weights were trained using only 496 data points
- Many unknowns with few data: overfitting

80% ± 5%

Output: prediction probability of positive IDH mutation

- Only single-value predictions; no uncertainty info

INTRODUCTION — Objective

- To develop computational capability for quantifying uncertainty in DNN models systematically and rigorously
- To determine the most sensitive factors in a model
- To assess robustness of model predictions against noise

METHODS

Model Setup

- Test data: MRI scans of 63 patients from The Cancer Genome Atlas (TCGA) database

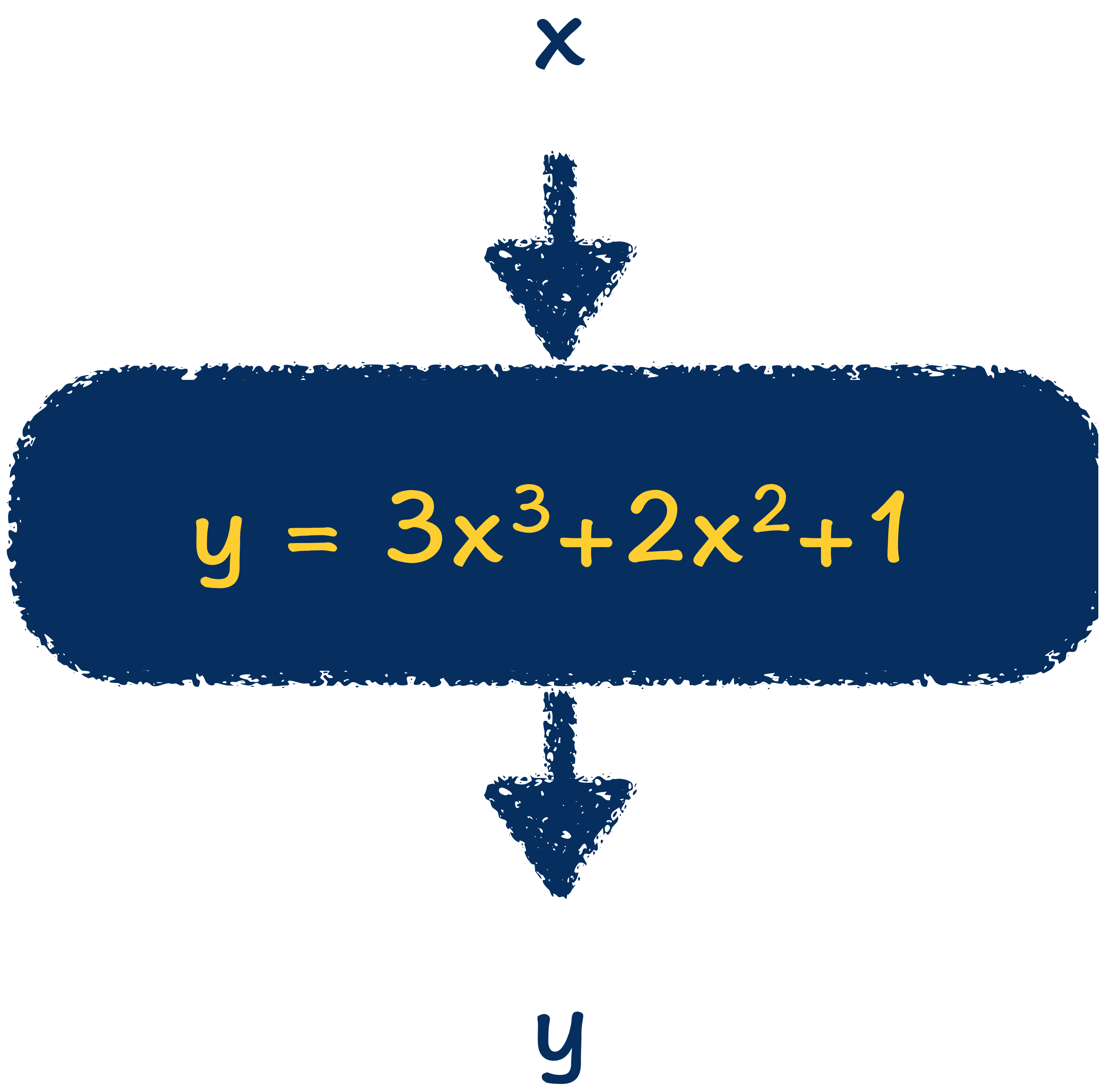
Noise in Model

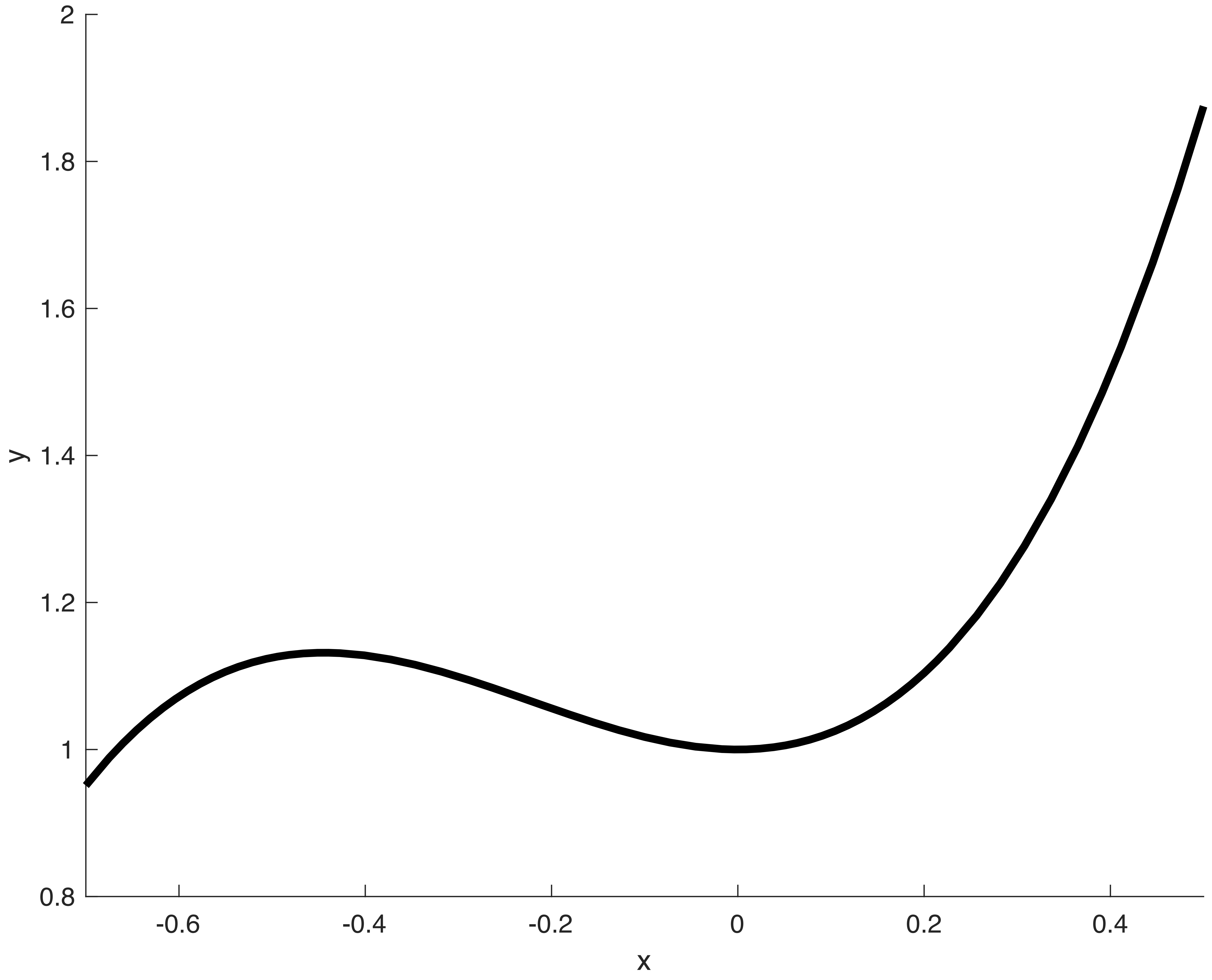
- Analyze the sensitivity of predictions with respect to trained model weights
- Add random Gaussian noise (1%, 5%, 10%) to the trained weights

Noise in Data

- Quantify uncertainty of output due to data
- Add random Rician noise to each pixel of MRI scans

METHODS



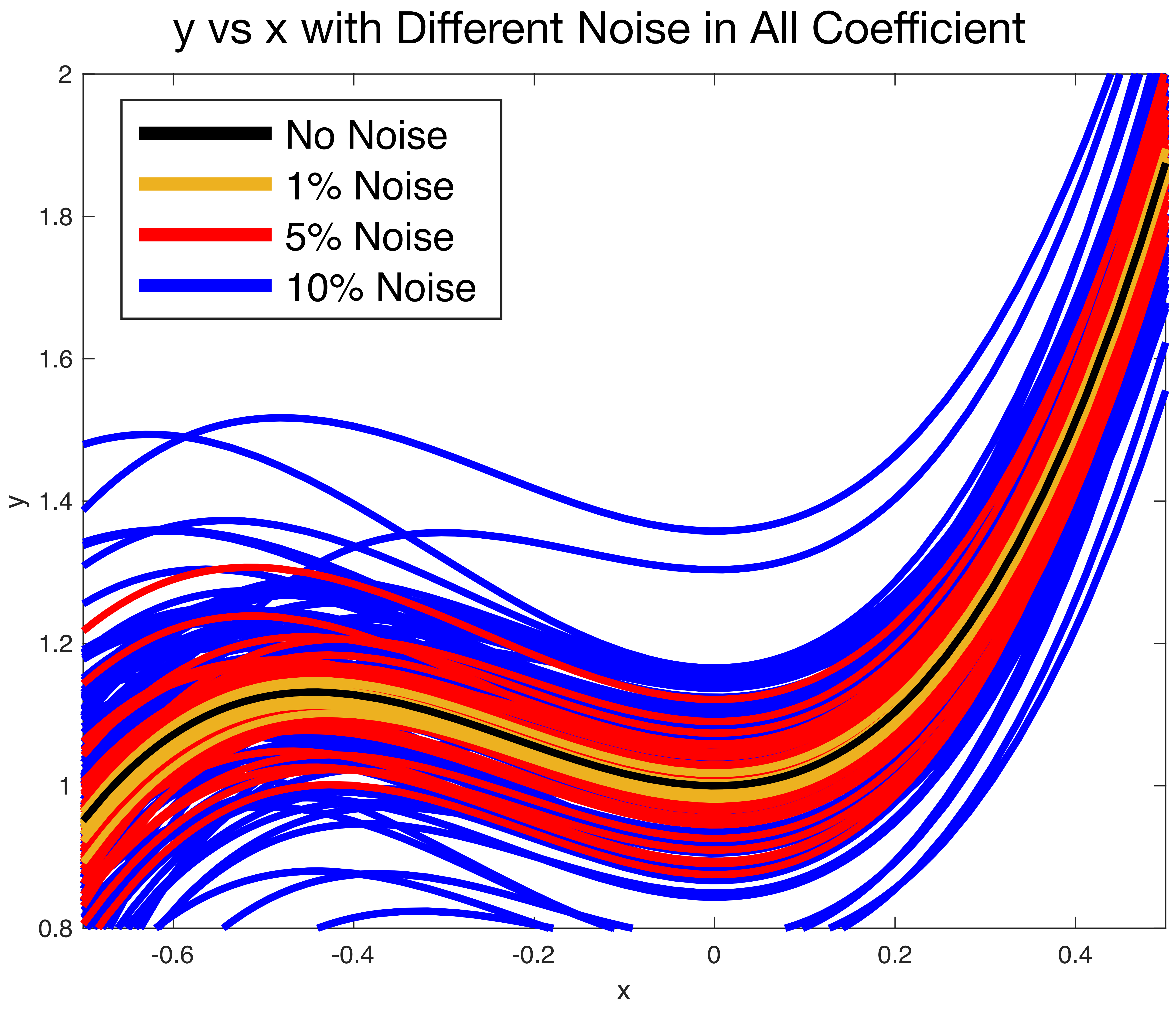


METHODS

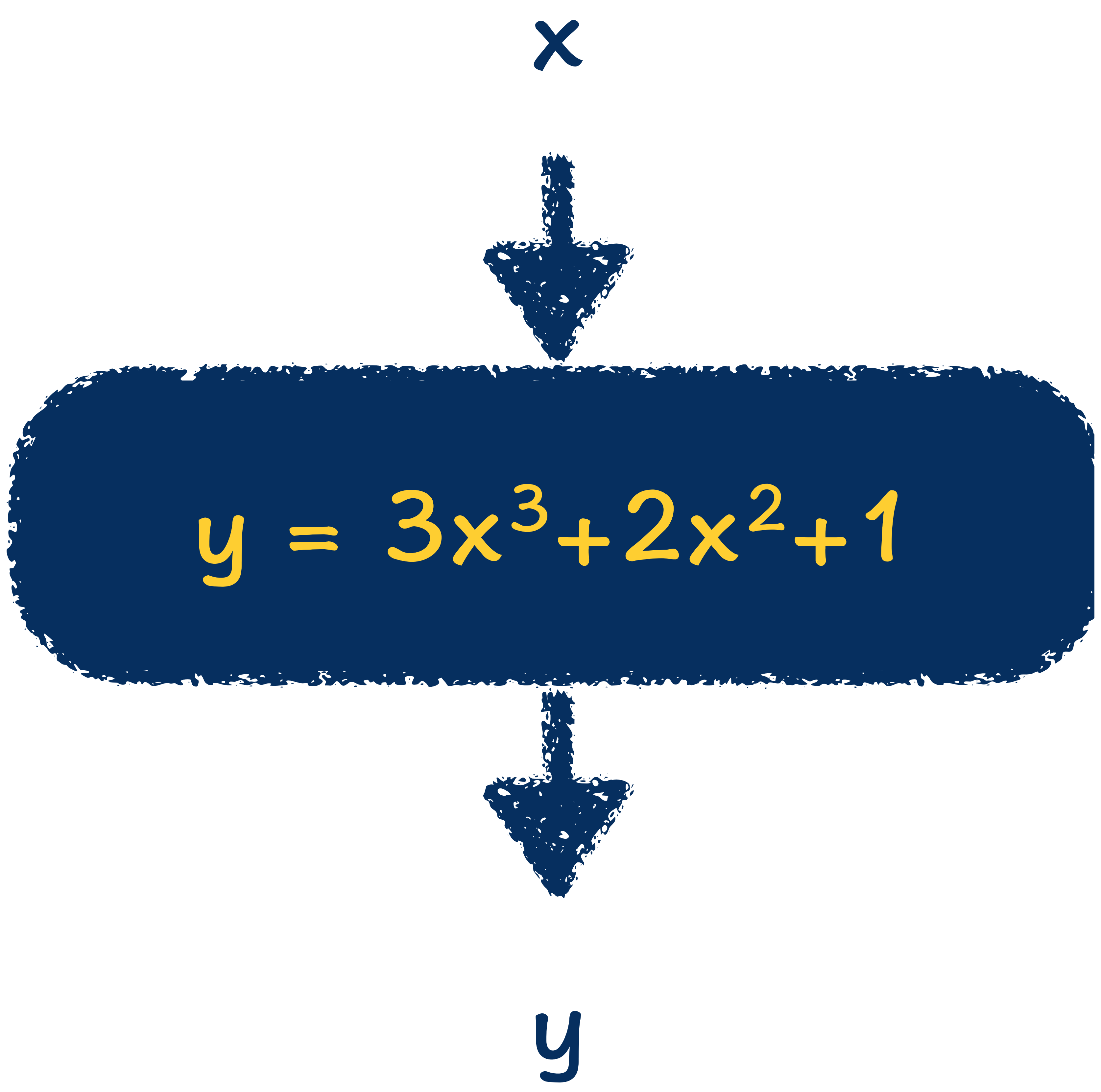
x

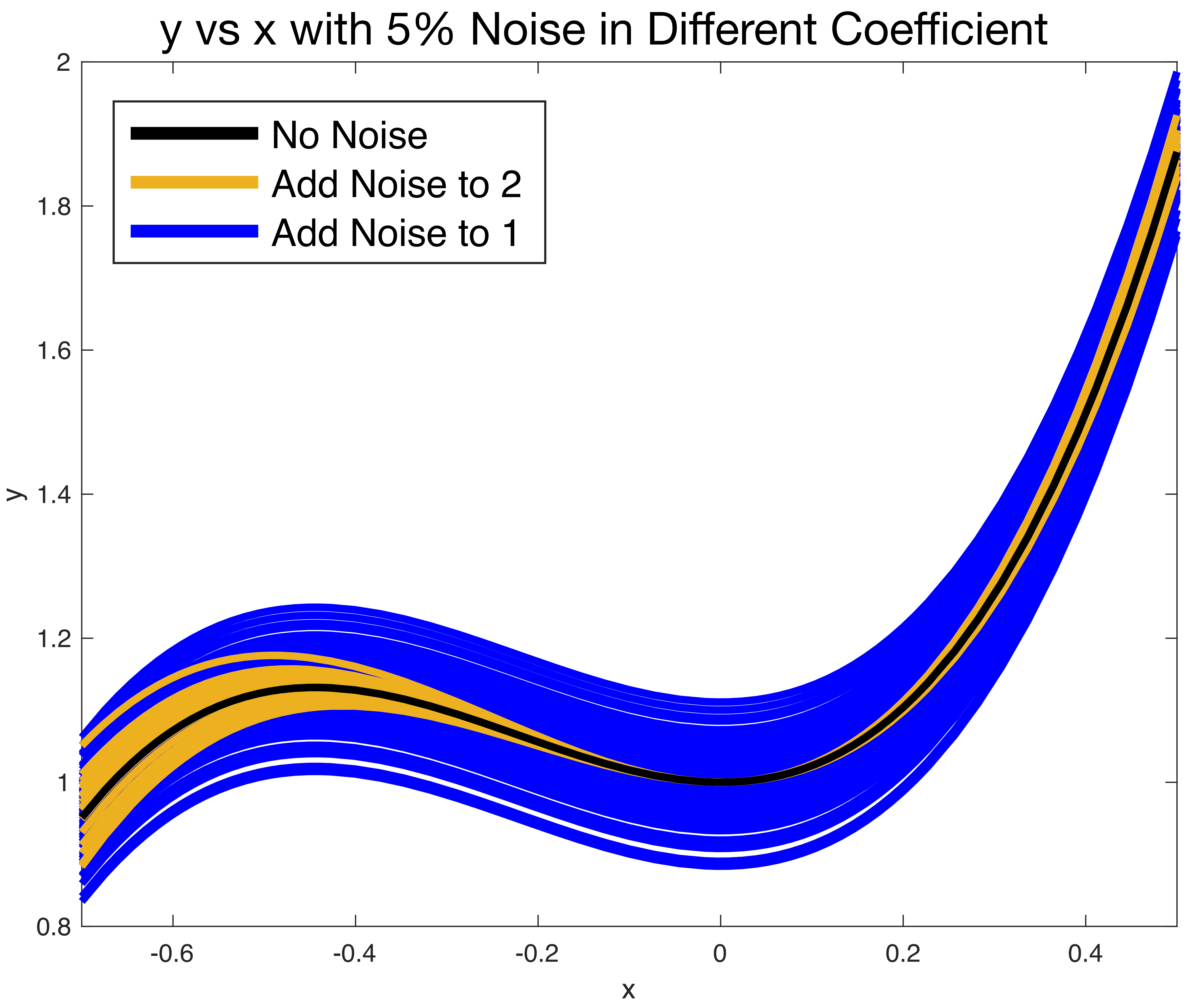
$$y = 3x^3+2x^2+1$$

y



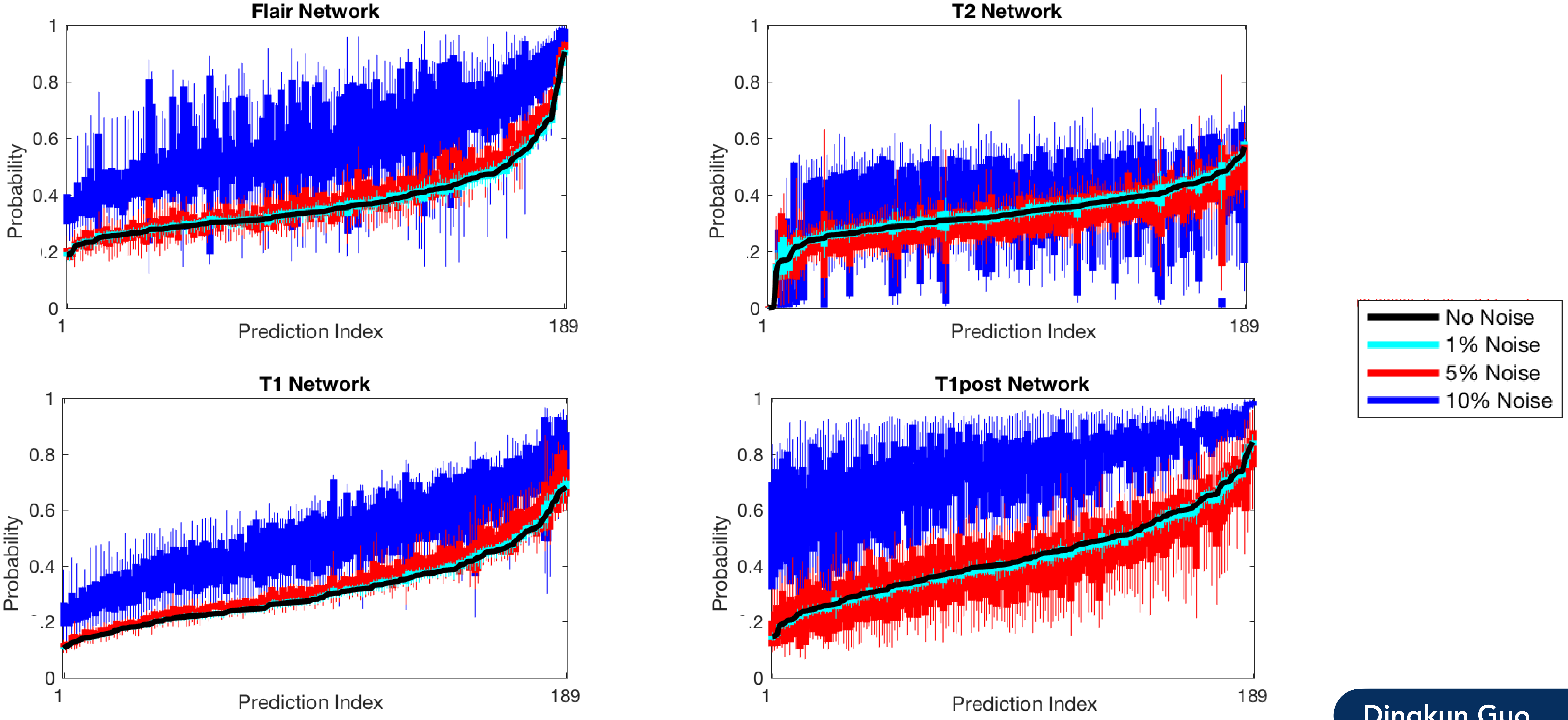
METHODS


x
$y = 3x^3+2x^2+1$
y


y vs x with 5% Noise in Different Coefficient
No Noise
Add Noise to 2
Add Noise to 1
y
x
2
1.8
1.6
1.4
1.2
1
0.8
-0.6
-0.4
-0.2
0
0.2
0.4

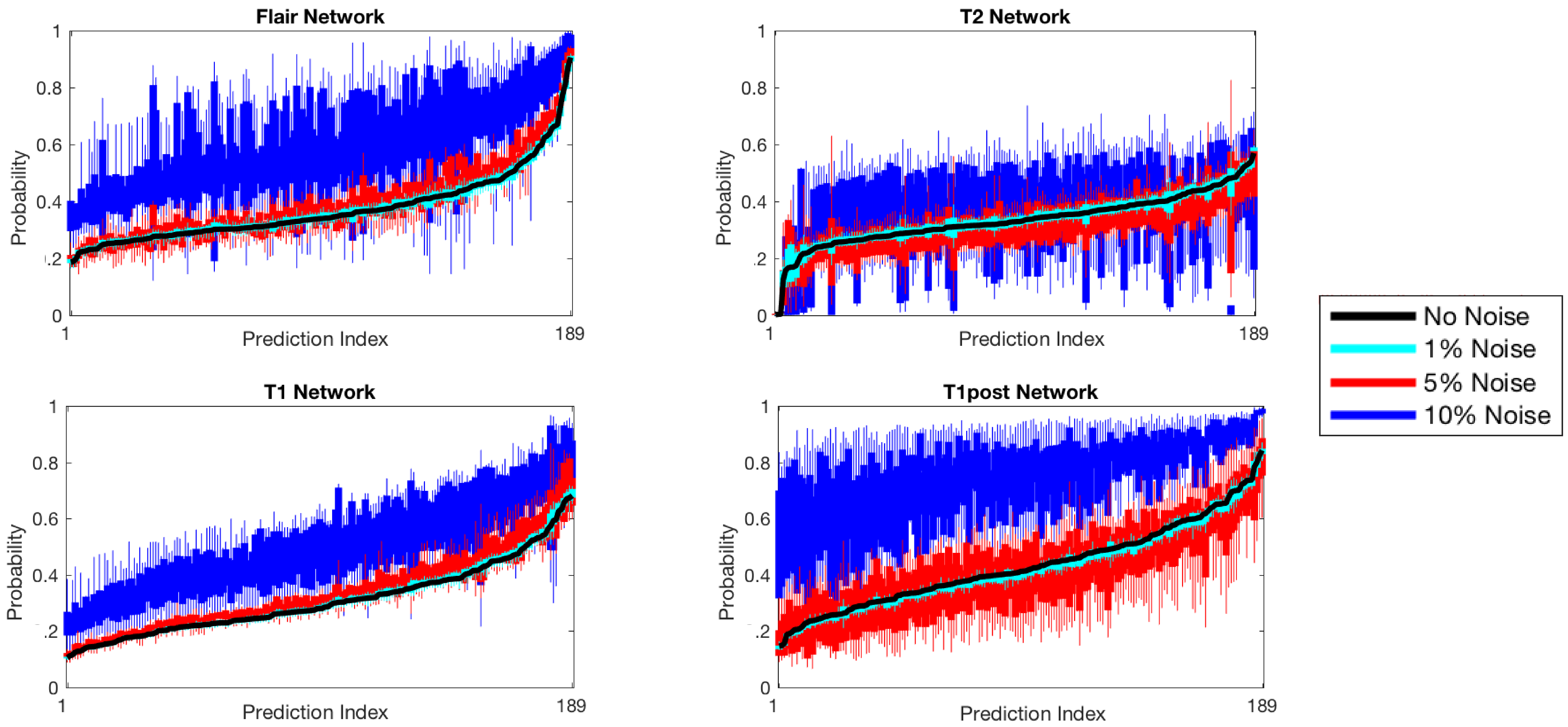
RESULTS — Noise in Model

- Adding Different Levels of Gaussian Noise to All Model Weights (10 Runs)



RESULTS — Noise in Model

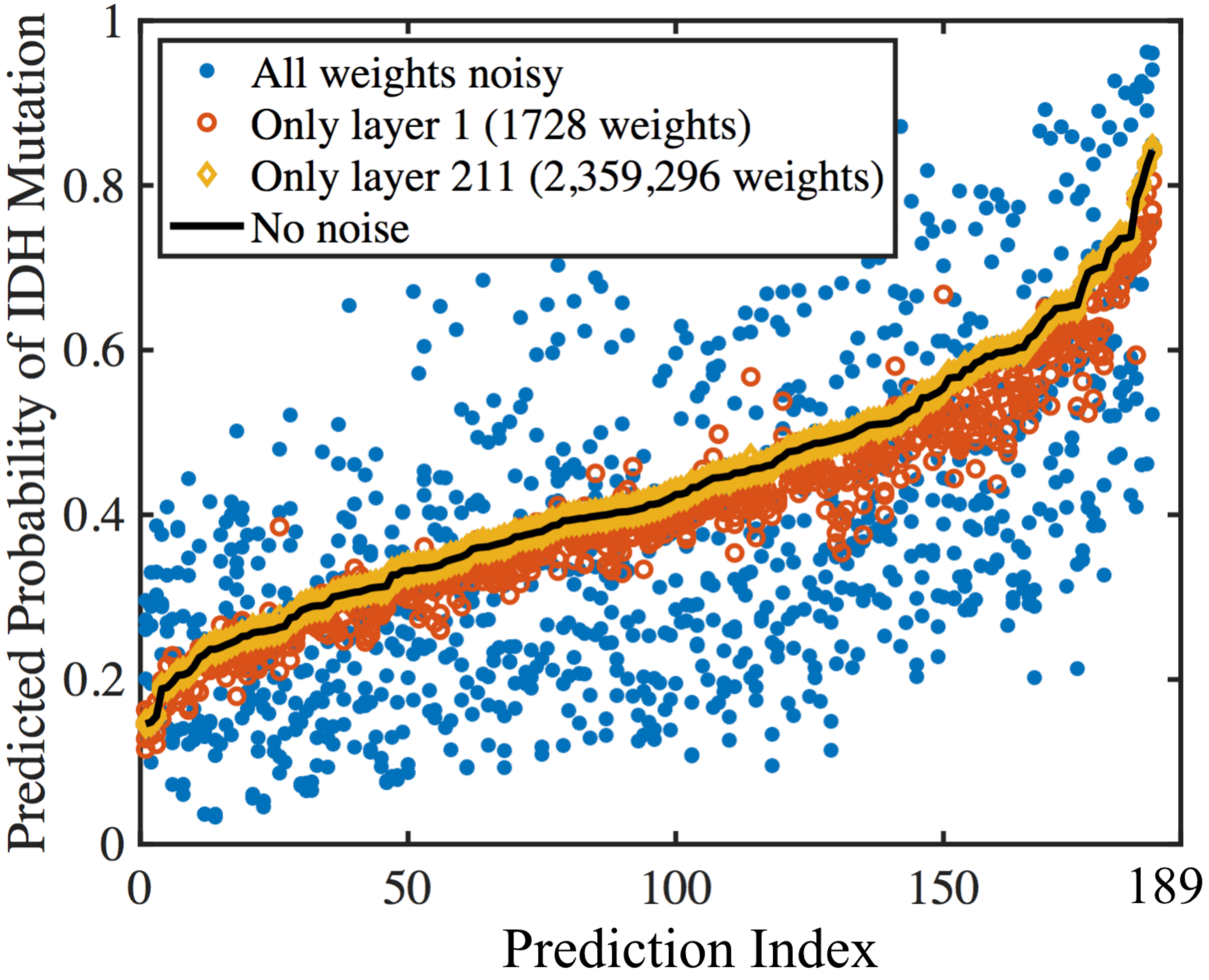
- **Adding Different Levels of Gaussian Noise to All Model Weights (10 Runs)**



- T1post Network is most sensitive to the noise
- Larger uncertainty observed when prediction probability is small
- T1 Network is less sensitive to large noise

RESULTS — Noise in Model

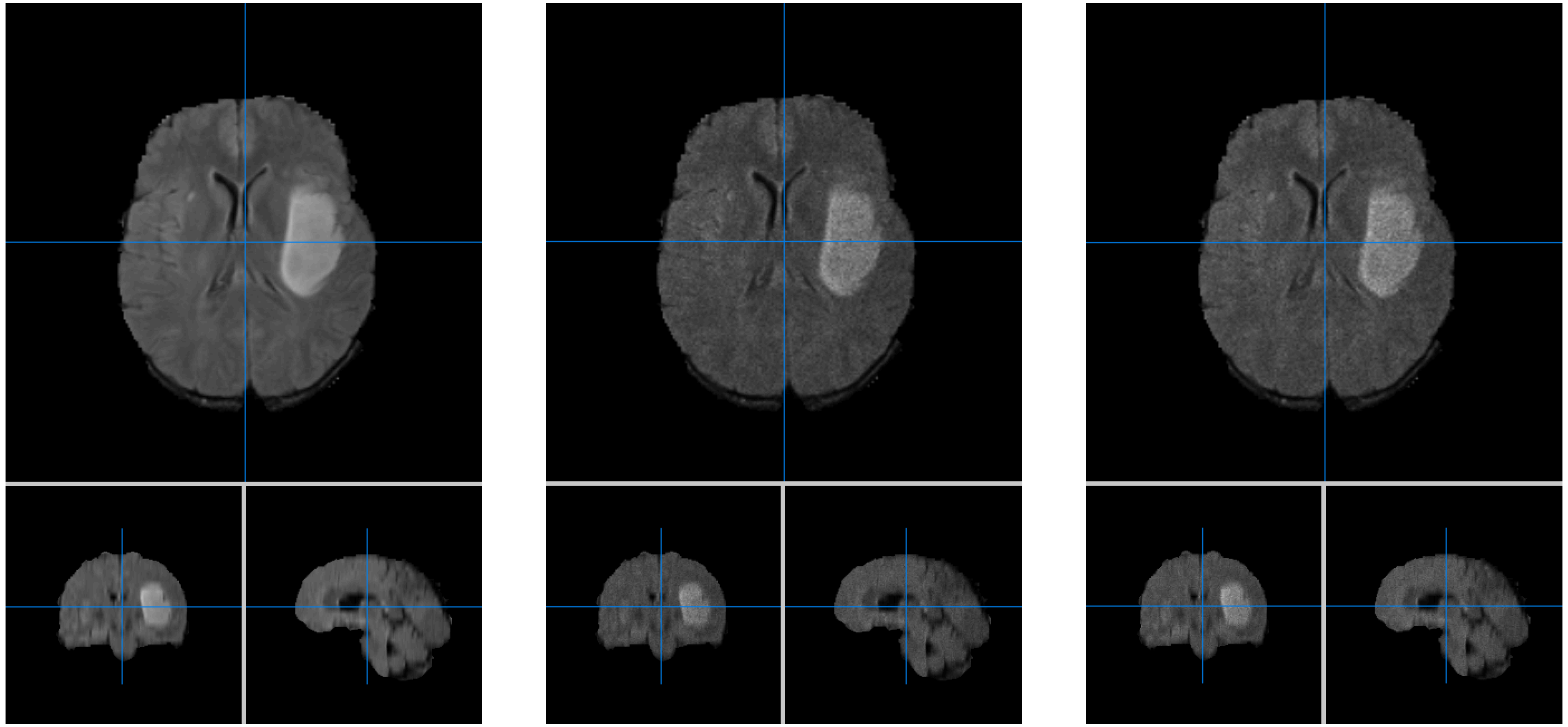
- **Adding Gaussian Noise to Weights in Different Layers**



- 5% Gaussian noise is added to different layers of T1post Network
- Although layer 211 has many more weights than layer 1, its noisy predictions have lower uncertainty
- Layer 1 contributes more to overall predictive uncertainty than layer 211
- This suggests some weights may be much more important than others

RESULTS — Noise in Data

- **Comparison of Different Types of Noise in MRI images**

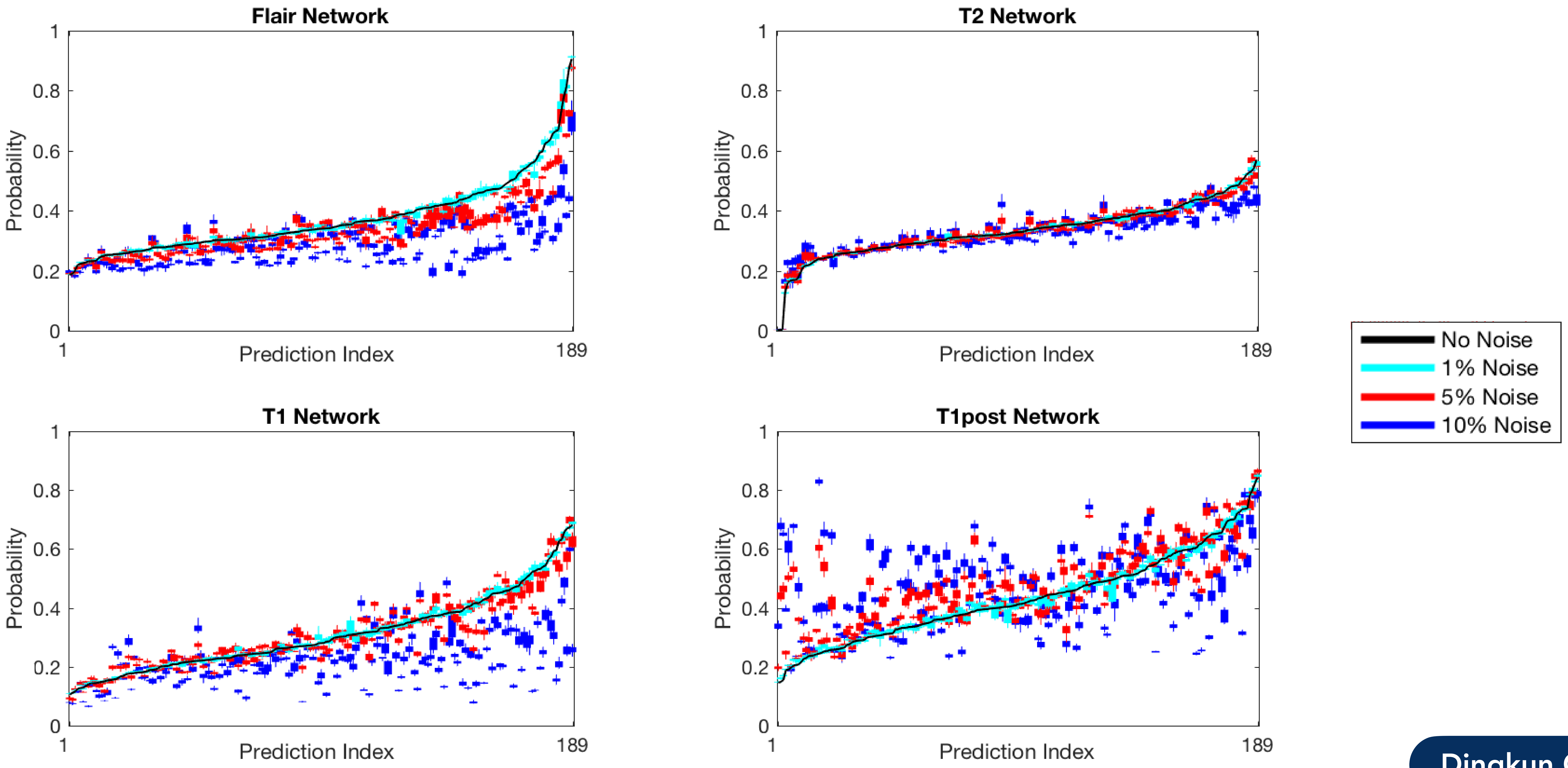


No Noise Image | 10% Gaussian Noise | 10% Rician Noise

- Different noise distributions in MRI images is hard to detect by human eyes, but can cause large difference in DNN predictions
- Rician noise is more commonly used in MRI applications [Gudbjartsson, 1995]

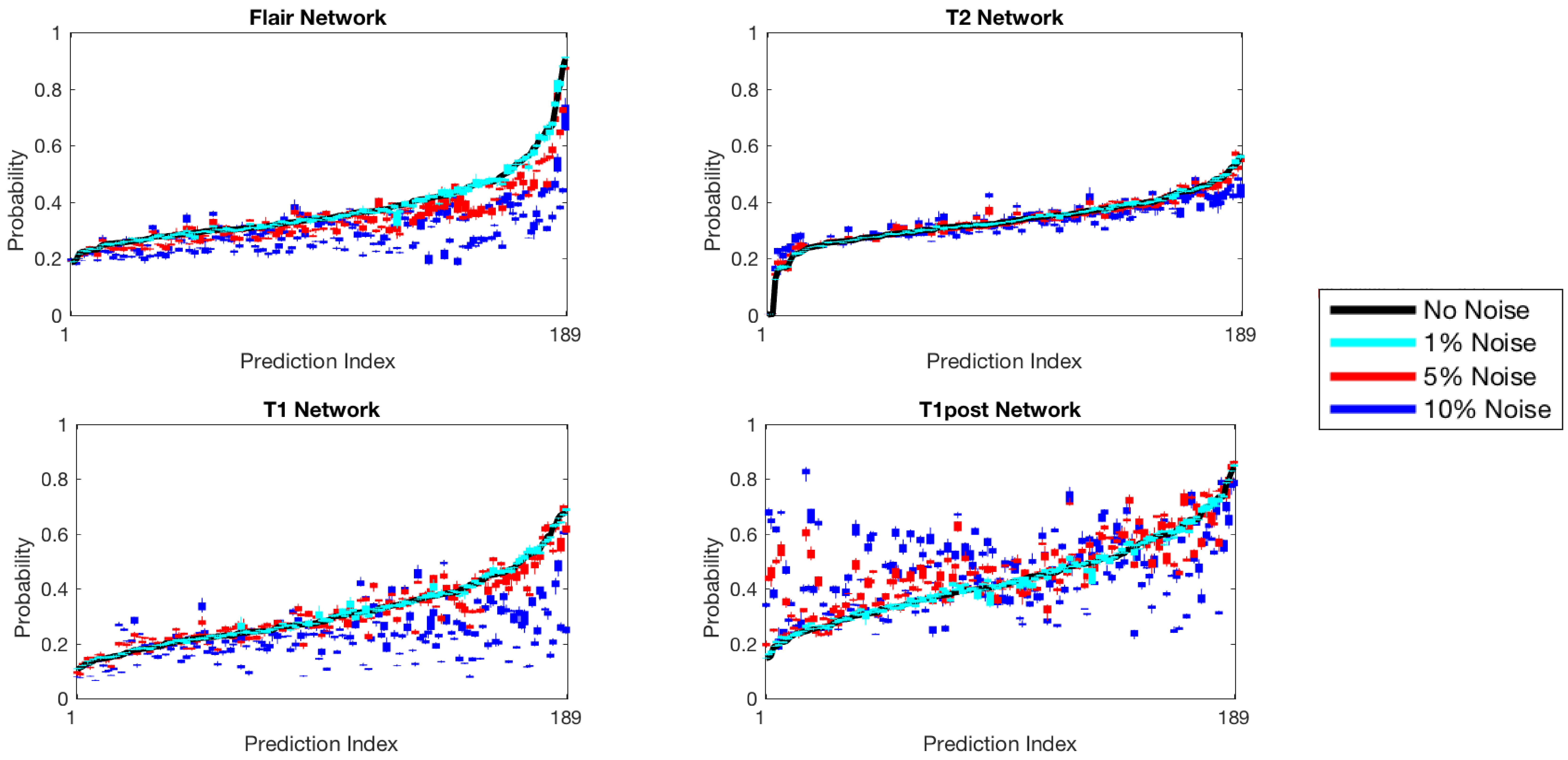
RESULTS — Noise in Data

- Adding Different Levels of Rician Noise to images (5 Runs)



RESULTS — Noise in Data

- **Adding Different Levels of Rician Noise to images (5 Runs)**



- Rician noise in images shifts predictions but produces less uncertainty than noise in weights
- The model appears more robust against image noise, but becomes less accurate

DISCUSSION

- Predictions can be quite sensitive to noise in model weights, which can be affected by the quality of training data and structure of the deep neural network.
- Uncertainty should be quantified and subsequently reduced from models and data
- Enabling high-confidence predictions is imperative for decision-making on patient treatments
- Uncertainty quantification will help hospitals and developers to compare and improve models

WHAT'S NEXT

- Analyze relationship between uncertainty and other features of tumor such as size/volume
- Conduct sensitivity analysis layer by layer
- Provide a list of criteria for assessing model robustness and generalizability
- Repeat model evaluations to produce additional data for statistical analysis

REFERENCE

- Chang et al. (2018) Residual Convolutional Neural Network for the Determination of IDH Status in Low- and High- Grade Gliomas from MR Imaging, Clin Cancer Res, 24(5).
- Pedano et al. (2016) Radiology Data from The Cancer Genome Atlas Low Grade Glioma [TCGA-LGG] collection. The Cancer Imaging Archive.
- Gudbjartsson and Patz (1995) The Rician Distribution of Noisy MRI Data. Magnetic Resonance in Medicine, 34(6), 910–914.

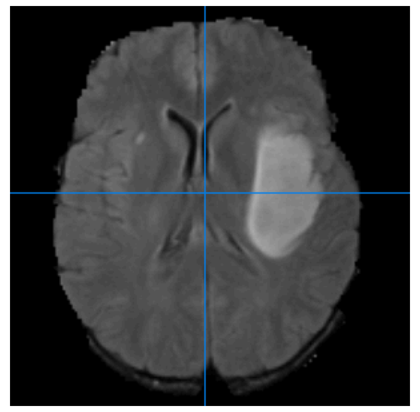
ACKNOWLEDGMENT

- Prof. Arvind Rao (University of Michigan Department of Computational Medicine & Bioinformatics and Department of Radiation Oncology)
- Srikanth Kuthuru and Nicholas Wang

Q & A

Q: Model Details?

Input: MRI scans of the brain



Model

- Step 1: registration and isotropic resampling
- Step 2: n4 bias correction and skull stripping
- Step 3: image intensity normalization
- Step 4: compile patient samples
- Step 5: prediction

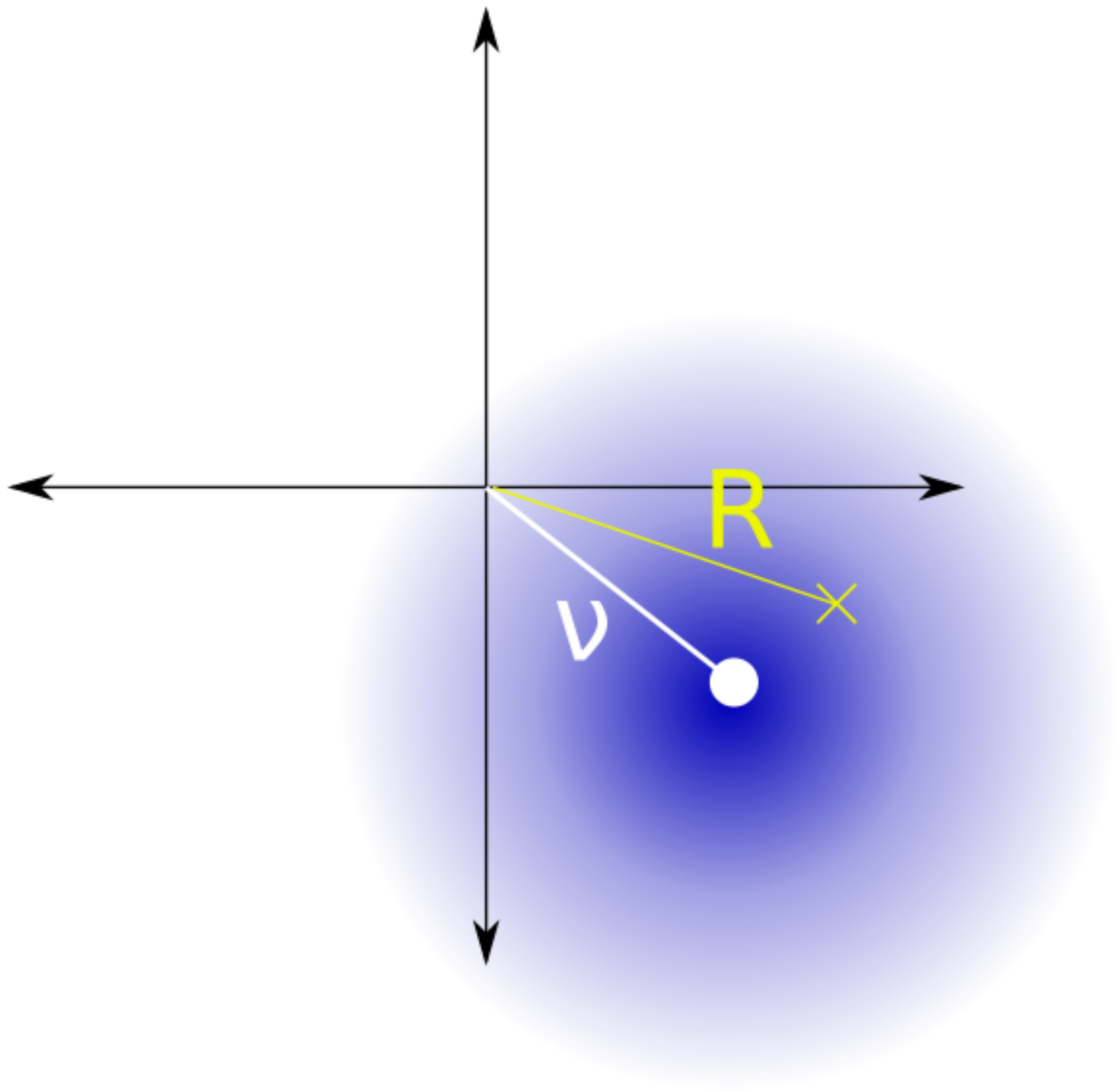
80%

Output: prediction probability of positive IDH mutation

Q: What is Rician Noise?

- $R = \sqrt{X^2 + Y^2}$ *where* $X \sim N(v\cos\theta, \sigma^2)$ *and* $Y \sim N(v\sin\theta, \sigma^2)$

- $R = \sqrt{X^2 + Y^2}$ *where* $X \sim N(v, \sigma^2)$ *and* $Y \sim N(0, \sigma^2)$



Thank you!