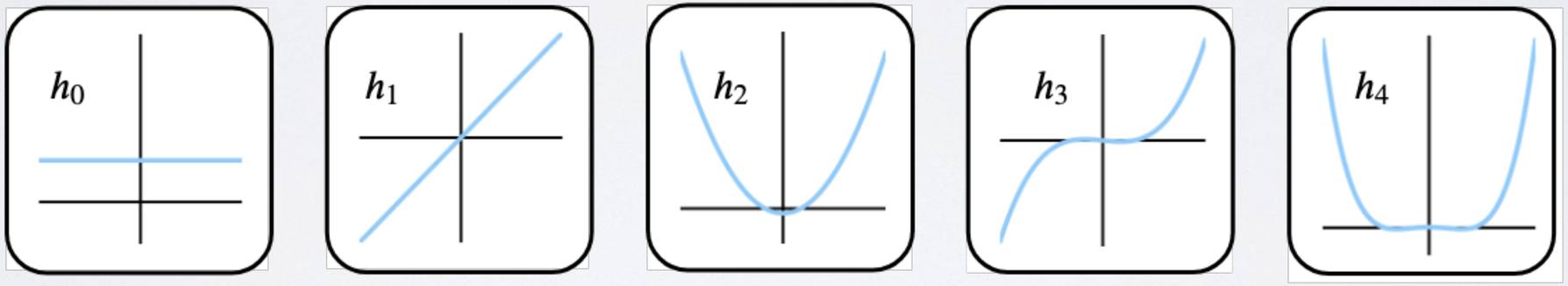
GENERATING ACCURATE PSEUDO-LABELS IN SEMI-SUPERVISED LEARNING AND AVOIDING OVERCONFIDENT PREDICTIONS VIA HERMITE POLYNOMIAL ACTIVATIONS






HERMITE POLYNOMIALS AS ACTIVATIONS

h_0 h_1 h_2 h_3 h_4

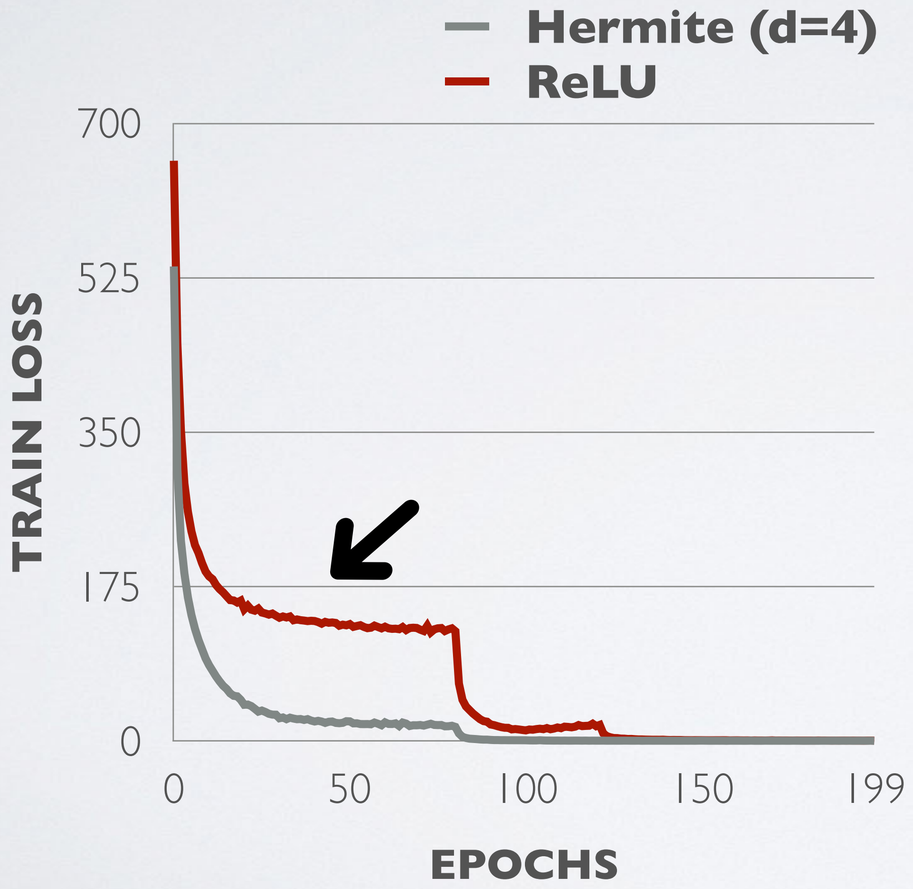


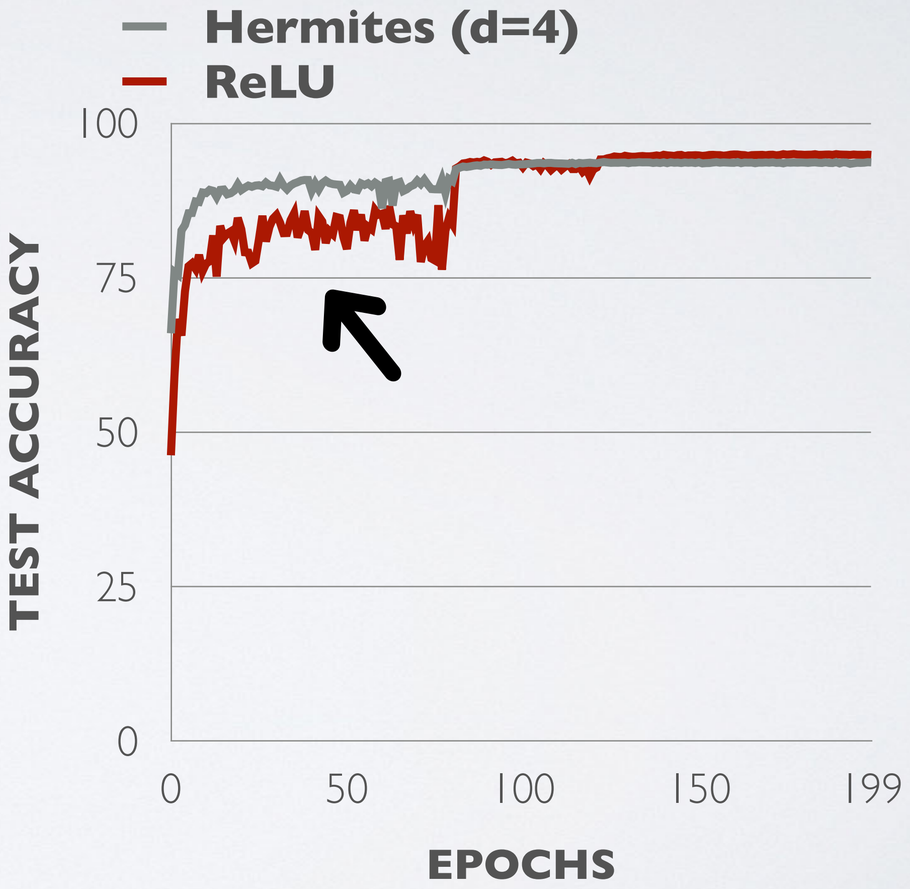
1 x $\frac{x^2-1}{2!}$ $\frac{x^3-3x}{3!}$ $\frac{x^4-6x^2+3}{4!}$

NON-SMOOTH RELUS TO POLYNOMIAL ACTIVATIONS

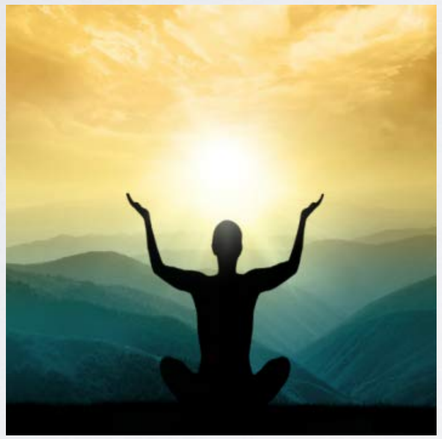
- Ge et al showed that we can avoid spurious local minima using Hermites
- Idea: use the lower order terms in the Hermite polynomial series **Expansion of ReLU** as **Activation Functions**

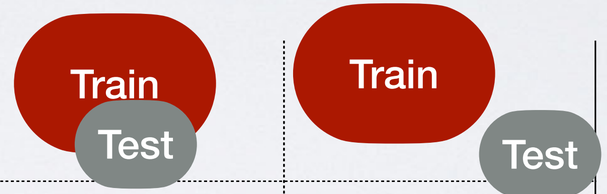
THE EARLY RISER PROPERTY IN HERMITES





HERMITES MAKE CONSCIOUS CLASSIFICATIONS



	Train **Similar** to Test	Train **Different** From Test	
			
ReLU	High Confidence Predictions	High Confidence Predictions	😕
Hermite	High Confidence Predictions	Low Confidence Predictions	😀

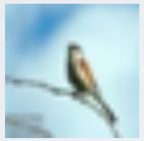
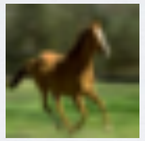
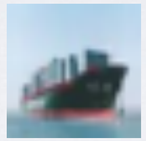
When the test data is different from the training data, Hermite networks consciously make (**approximately) random** predictions unlike ReLU networks

THE SAAS CONCEPT

PSEUDO-LABELS GENERATION VIA THE BANDERSNATCH PHENOMENON

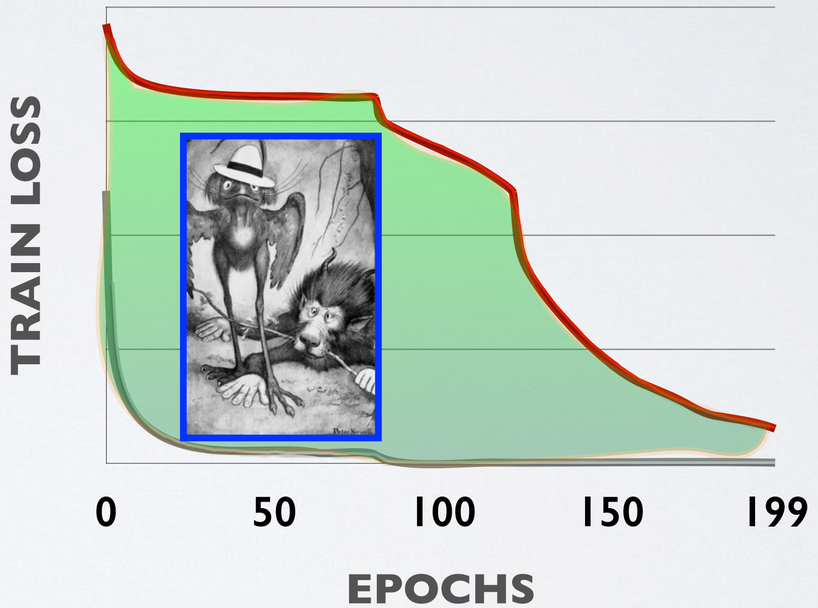
Good Labels ➡ **Least Training Time**

CIFAR10

TRUE LABELS Bird Horse Ship

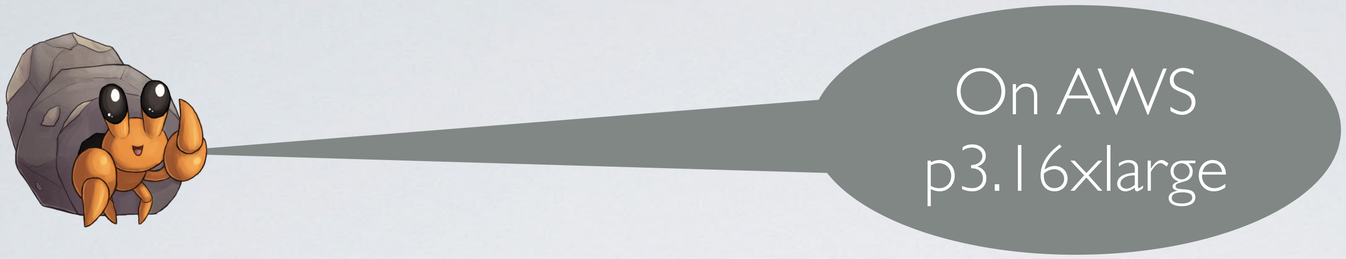
RANDOM LABELS Cat Truck Dog



SaaS: Find labels with least training time

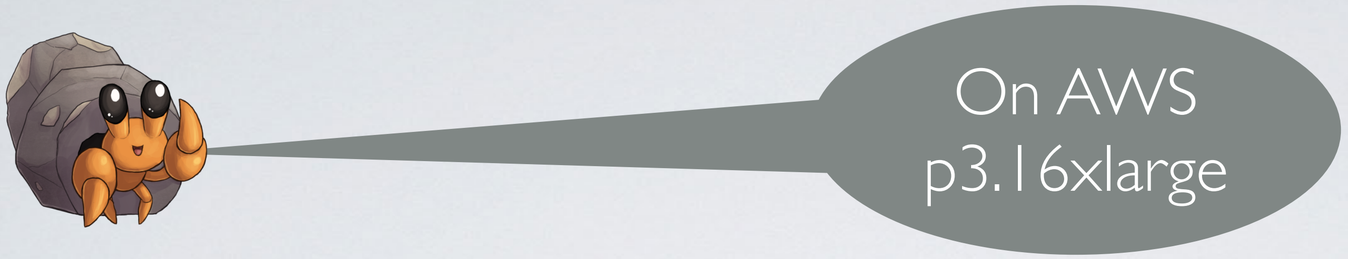
*bandersnatch ~ a fast and intelligent fictional creature https://en.m.wikipedia.org/wiki/Bandersnatch

COMPUTATIONAL BENEFITS



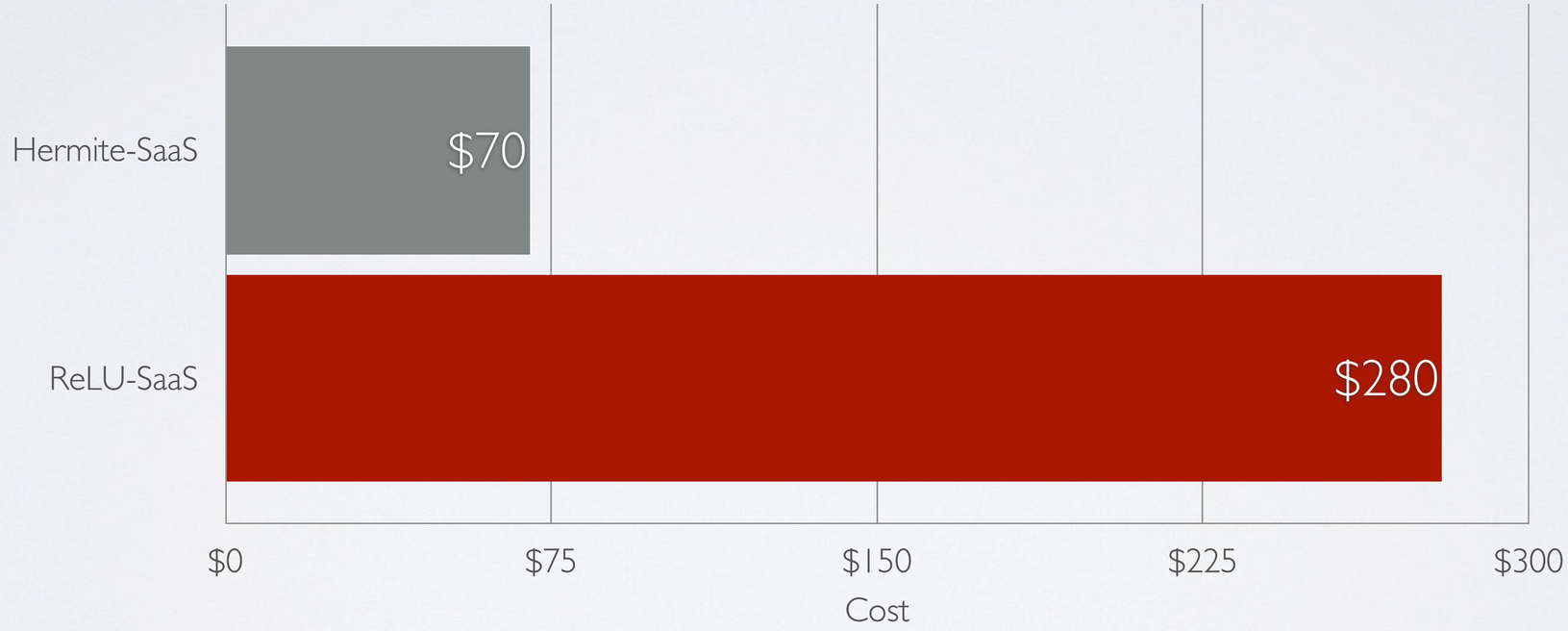
HERMITE-SAAS TRAINS FASTER THAN RELU-SAAS

CIFAR10



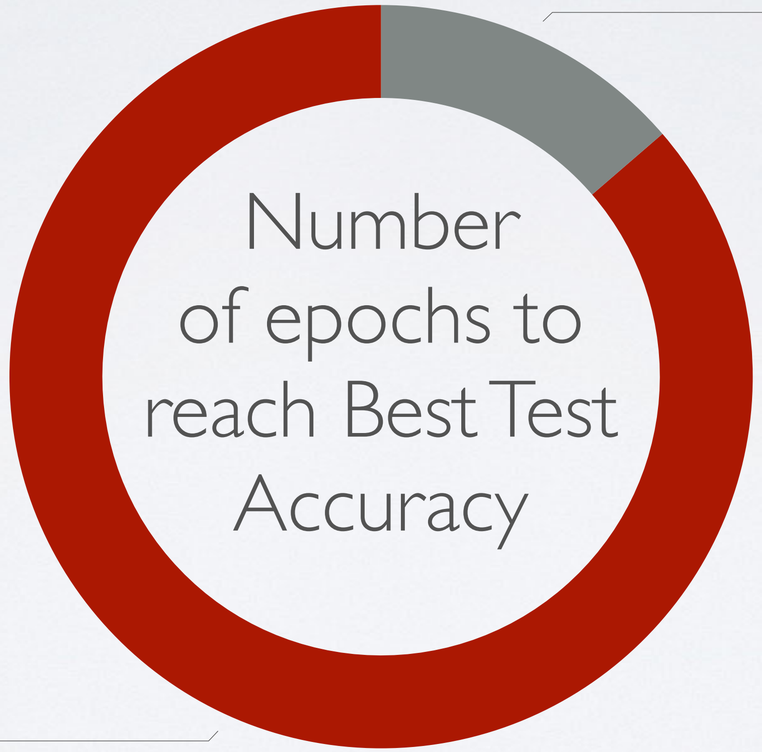
HERMITE-SAAS COSTS LESS THAN RELU-SAAS

CIFAR10

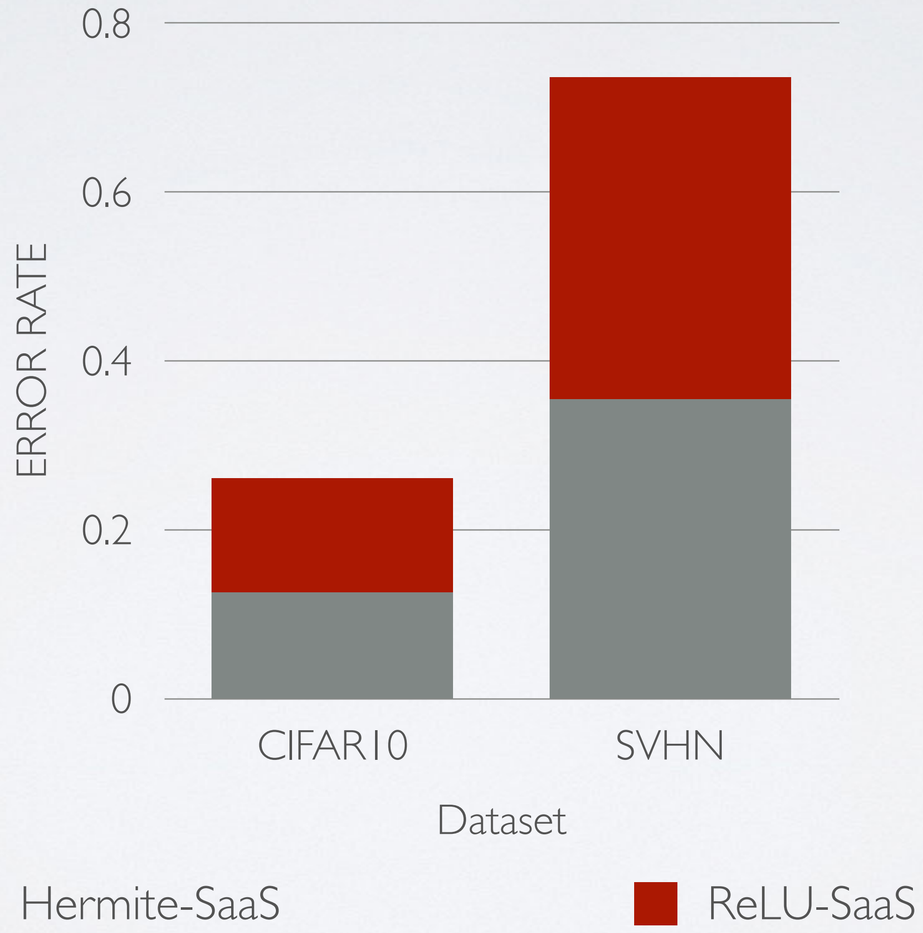


Adding 30% uniform noise to labelled data

HERMITE-SAAS MORE RESILIENT TO NOISE THAN RELU-SAAS



HERMITE-SAAS GENERALIZES BETTER THAN RELU-SAAS




Hermite polynomials
are nice