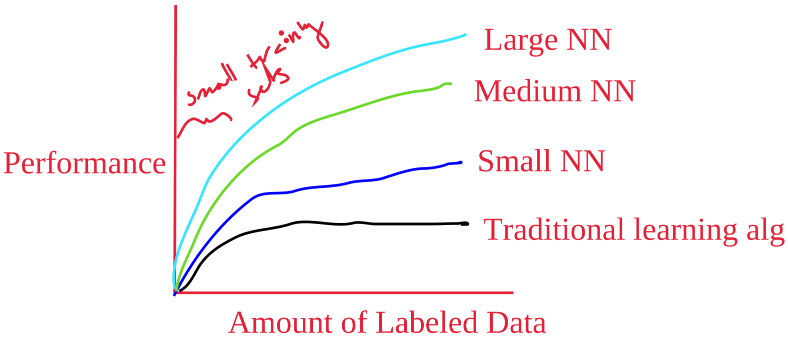
Audio & translators -> RNN or Custom

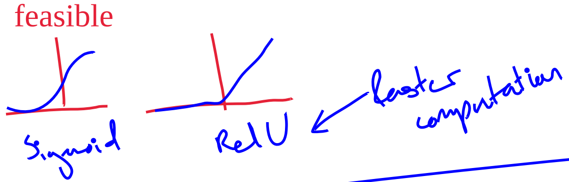
Audio is considered Unstructured data

V3



As Data and NN size increases, time to train also increases

As data and computation capacity increases, NN becomes more feasible



V4

Binary Classification

$(x, y) \quad x \in \mathbb{R}^{n_x},\ y \in \{0, 1\}$

Where x is a n-dimensional feature vector and y can be 0 or 1

m training examples : $\{(x^{(1)}, y^{(1)}) \ldots (x^{(m)}, y^{(m)})\}$

m = m_train

m_test = #test examples

$$X = \underbrace{\begin{bmatrix} | & | & & | \\ x^{(1)} & x^{(2)} & \cdots & x^{(m)} \\ | & | & & | \end{bmatrix}}_{m} \Big\updownarrow n \qquad Y = [y^{(1)}, y^{(2)} \ldots y^{(m)}]$$

$X \in \mathbb{R}^{n \times m} \qquad X.\text{shape} = (n, m)$

$Y \in \mathbb{R}^{1 \times m}$

$Y.\text{shape} = (1, m)$

V7

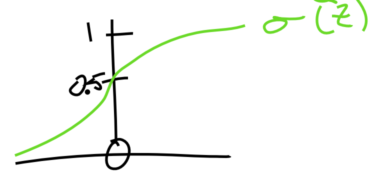
Logistic Regression

Given x, want $\hat{y} = P(y=1 \mid x)$

probability that y is 1 given x

$x \in \mathbb{R}^n \qquad$ Parameters: $w \in \mathbb{R}^n,\ b \in \mathbb{R}$

Output $\hat{y} = \sigma(\underbrace{w^T x + b}_{z}) \qquad \sigma(z) = \frac{1}{1+e^{-z}}$



If z large $\sigma(z) \approx \frac{1}{1+0} = 1$

If z large negative #

$\sigma(z) = \frac{1}{1+e^{-z}} \approx \frac{1}{1+\infty} \approx 0$

V8

Logistic Regression cost function

$\hat{y} = \sigma(w^T x + b)$, where $\sigma(z) = \frac{1}{1+e^{-z}}$

Given $\{(x^{(1)}, y^{(1)}) \ldots (x^{(m)}, y^{(m)})\}$

want $\hat{y}^{(i)} \approx y^{(i)}$

Lose (error) Function: $\mathcal{L}(\hat{y}, y) = \frac{1}{2}(\hat{y} - y)^2$

$$\mathcal{L}(\hat{y}, y) = -(y \log \hat{y} + (1-y)\log(1-\hat{y}))$$

If $y=1$: $\mathcal{L}(\hat{y}, y) = -\log \hat{y} \leftarrow$ want $\log \hat{y}$ large

If $y=0$: $\mathcal{L}(\hat{y}, y) = -\log(1-\hat{y}) \leftarrow$ want $\log 1-\hat{y}$ large, want $\hat{y}$ large, want $\hat{y}$ small

Cost function: $J(w, b) = \frac{1}{m}\sum_{i=1}^{m} \mathcal{L}(\hat{y}^{(i)}, y^{(i)})$

$$= -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)} \log \hat{y}^{(i)} + (1-y^{(i)})\log(1-\hat{y}^{(i)})\right]$$

V9

$\hat{y} = \sigma(w^T x + b),\ \sigma(z) = \frac{1}{1+e^{-z}}$

$J(w,b) = \frac{1}{m}\sum_{i=1}^{m} \mathcal{L}(\hat{y}^{(i)}, y^{(i)}) = -\frac{1}{m}\sum_{i=1}^{m} y^{(i)} \log \hat{y}^{(i)} + (1-y^{(i)})\log(1-\hat{y}^{(i)})$

Want to find w,b that minimize J(w,b)