

# Machine learning problem formulation

- The machine learning approach expresses NLP as an optimization problem:

$$\hat{\mathbf{Y}} = \operatorname{argmax}_{\mathbf{y} \in f(\mathbf{x}; \boldsymbol{\theta})} \Psi(\mathbf{Y}, \mathbf{X}; \boldsymbol{\theta})$$

*where*  $\mathbf{x} \in X$  is the input

$\mathbf{y} \in Y$  is the output

$\Psi(\cdot) \rightarrow \mathbb{R}$  is a function expressing the learning objective

$f(\cdot)$  is the function, or model, that maps  $\mathbf{x}$  to  $\mathbf{y}$

$\boldsymbol{\theta}$  parameterizes  $f(\cdot)$

# Neural Language Modeling

- End-to-end modeling with no explicit syntactic representation
  - Neural networks pretrained on very large text corpora where goal is to predict words given then context (i.e., surrounding words).
  - Transfer / few shot learning to fine tune on downstream tasks

