

Sequence Labeling (Tagging), Tagging with Classifiers

Input $\bar{x} = (x_1, x_2, \dots, x_n)$

Output $\bar{y} = (y_1, y_2, \dots, y_n)$ one prediction per word

Structured classification Fed raises interest rates 0.5 percent

Predict each y_i independently w/ logistic regression

$P(y_i = y | \bar{x}, i)$ index we're predicting at
 $f(\bar{x}) = [\underset{\text{Fed}}{1} \ 0 \ 0 \ \underset{\text{raises...}}{1} \ 0 \ 0 \ 0 \ \dots]$

BOW: $f(\bar{x}, y = \text{NN}) = [\frac{000f(\bar{x})000}{\text{NN}} \mid \frac{100100}{\text{NN}} \mid \frac{0000}{\text{NN}} \dots]$

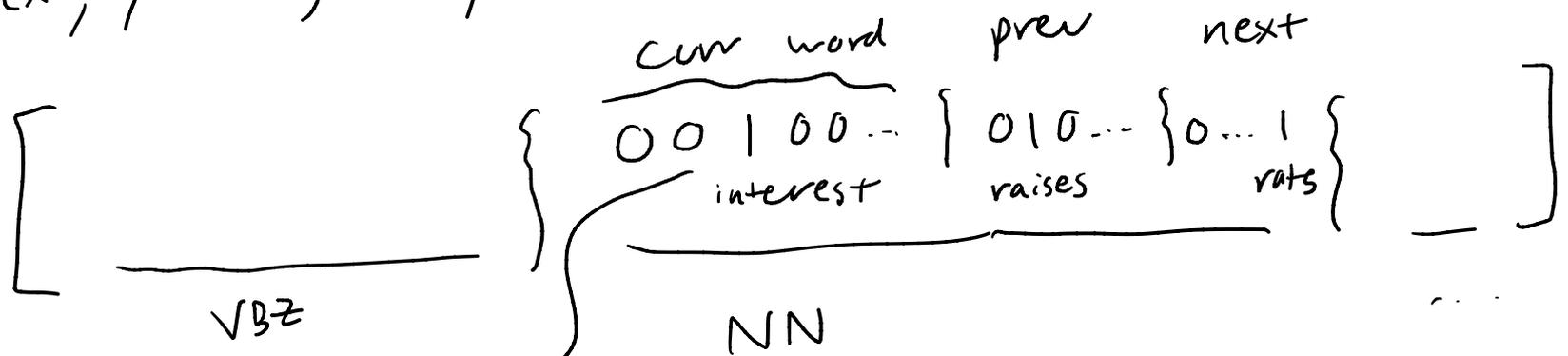
independent of i

Simple way: single feature on curr word

Positional: $f(\bar{x}, y = \text{NN}, i=3) [0 \text{---} \mid \frac{00100000}{\text{interest NN}} \mid 0 \text{---}]$

Positional features with context

$f(x, y = NN, i = 3)$



Conjunction of properties

Indicator ["cur word = interest \wedge tag = NN"]

"word" in a bag-of-words space

Problem w/ classification for tagging

indicators \rightarrow classifier

What goes wrong?

VBZ \rightarrow ~~VBP~~

NNS \rightarrow ~~NN~~

Fed raises interest rates ...

predictions of a classifier may be incoherent

Sequence modeling:

Hidden Markov Models

Conditional Random Fields