bigrams		encoded		onehot X	weights W	Y predicted
Χ	Υ	Χ	Υ			
\Box	е	0	5	100000000000000	27,27	
e	m	5	13	000001000000000	27x27 random	nBigrams x 27
m	m	13	13	00000000000010	random	
m	a	13	1	00000000000010		
a	.	1	0	010000000000000		
	0	0	15	000000000000001.		

Forward Pass

$$Yp = W * X + b$$

$$Loss = 1/n \sum 1/2(Yp - Y)^2$$

Backward Pass

$$\begin{split} L &= \frac{1}{2} (Yp - Y)^2 &= \frac{1}{2} ((Yp)^2 - 2YpY + Y^2) &= (Yp)^2/2 - YpY + Y^2/2 \\ &\frac{\partial L}{\partial Yp} = Yp - Y \\ &\frac{\partial L}{\partial W} = \frac{\partial L}{\partial Yp} * \frac{\partial Yp}{\partial W} \quad \text{and} \quad \frac{\partial Yp}{\partial W} = X \\ &\text{SO} \quad \frac{\partial L}{\partial W} = \frac{1}{n} \sum (Yp - Y) * X \end{split}$$