

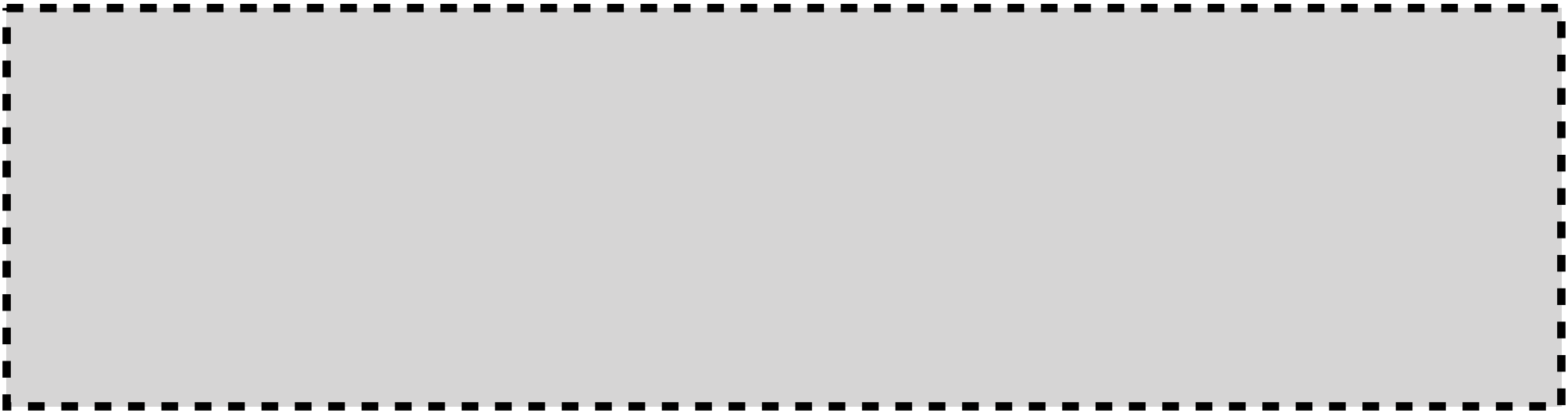
Dual RNN LM





Output Combination : Trainable switch parameter











V

1

V2





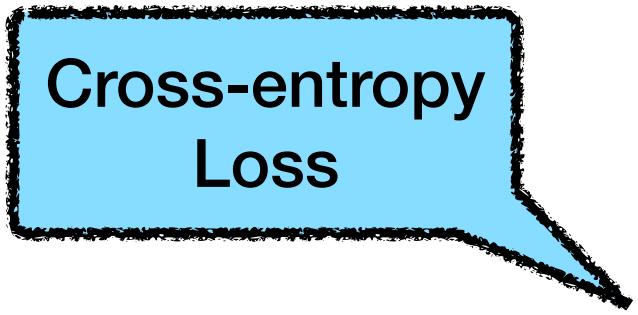
I

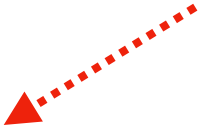
O

S

S

Cross-entropy Loss







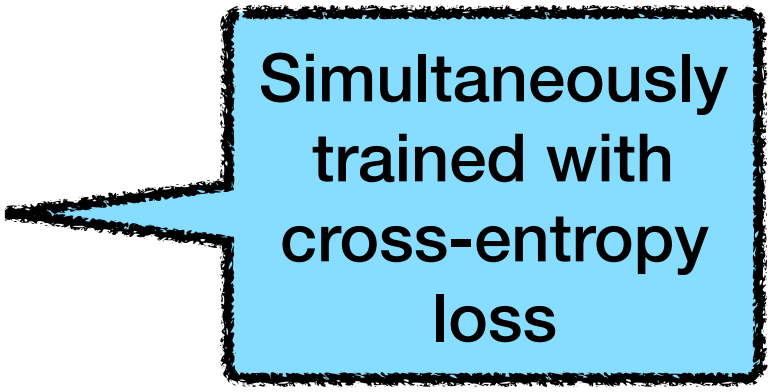
Total Loss

p



A diagram illustrating the decomposition of a total loss. At the top center, the text "Total Loss" is written in a large, black, italicized serif font. Below this, two black arrows point upwards towards the text. The left arrow originates from the bottom left and points towards the left side of "Total Loss". To its left, the red italicized symbol p is positioned. The right arrow originates from the bottom right and points towards the right side of "Total Loss". To its right, the red italicized expression $1-p$ is positioned.

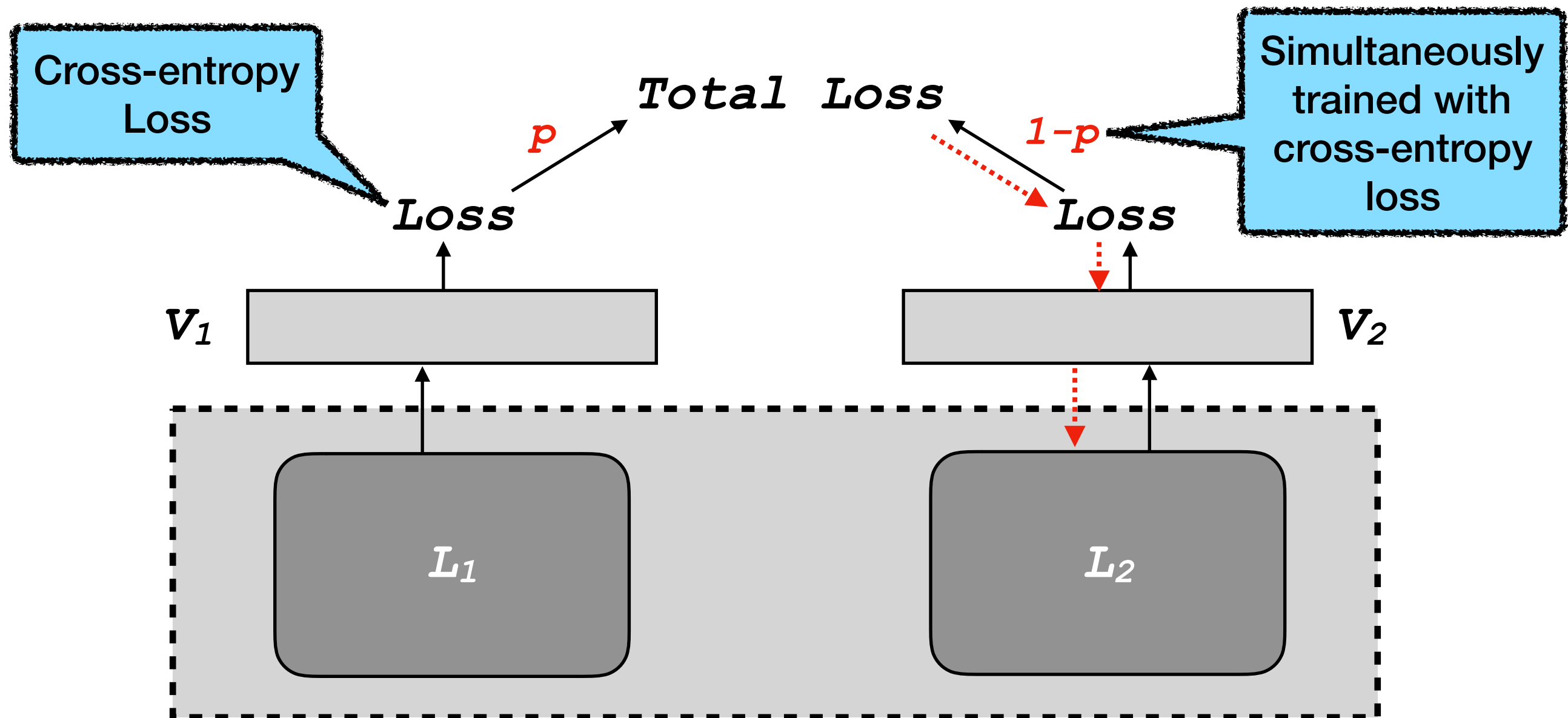
$1-p$



**Simultaneously
trained with
cross-entropy
loss**

Dual RNN LM

Output Combination : Trainable switch parameter



Results

Perplexity	Data	RNNLM	DLM
No pre-training	Dev	94.1	93.65
	Test	78.8	75.49