

Author: Benjamin Smidt  
Created: September 22, 2022  
Last Updated: September 22, 2022

# Deep Learning Resources

## Contents

<b>1</b>	<b>General</b>	<b>1</b>
<b>2</b>	<b>Data Preprocessing</b>	<b>1</b>
2.1	Weight Initialization . . . . .	1
<b>3</b>	<b>Loss Functions</b>	<b>1</b>
<b>4</b>	<b>Architectures</b>	<b>2</b>
<b>5</b>	<b>Hyperparameters</b>	<b>2</b>
5.1	Update Rules . . . . .	2
5.2	Regularization . . . . .	2
5.3	Dropout . . . . .	2

## 1 General

## 2 Data Preprocessing

### 2.1 Weight Initialization

## 3 Loss Functions

1. Mikolov et al., “Distributed Representations of Words and Phrases and their Compositionality”, ArXiv 2013 [*Hierarchical Softmax (Problem of Many Classes)*]

## 4 Architectures

## 5 Hyperparameters

### 5.1 Update Rules

### 5.2 Regularization

### 5.3 Dropout