NLP HW4-5 Report

tags: `class` `uva`

1. Word Embeddings

1.1

Training accuracy = 0.515975

Dev accuracy = 0.5322

1.2

Training accuracy = 0.641750

Dev accuracy = 0.6314

1.3

Training accuracy = 0.638350

Dev accuracy = 0.6328

2. Recurrent Neural Network Language Models

2.2

Perplexity on training data: 487.7070748995684

Perplexity on dev set: 462.1976028417219

2.3

Best n: 3

Perplexity with n=3 on training data: 481.2340465363492

Perplexity with n=3 on dev set: 423.11754206546567

2.4

optimization method: Adagrad

Perplexity with adaGrad on dev: 391.6149915554

perplexity with adaGrad on train: 378.6199790127475

2.5

Best Hiddem dim & embedding

hidden dim 256

embedding 256

dev perplexity: 296.51765534800467

train perplexity: 170.4561469277748

2.6

Yes. Different batch sizes do make a difference.

best batch size: 16

dev: 689.4363125426024

train: 897.1313238734933