

A Experimental setup

Initialization for text classification. Hidden-to-hidden weight matrices W^h are initialized orthogonally and all other matrices are initialized uniformly using the method by [Glorot and Bengio \(2010\)](#).

We train our networks using batches of size 128 and a learning rate of 0.0005.

Initialization for language modeling. All weight matrices of the networks are initialized orthogonally and all biases are initialized with zeros. Initial values of hidden and cell elements are not trainable and equal to zero.

For the character-level task, we train our networks on non-overlapping sequences of 100 characters in mini-batches of 64 using a learning rate of 0.002 and clip gradients with threshold 1.

For the word-level task, networks are unrolled for 35 steps. We use the final hidden states of the current mini-batch as the initial hidden state of the subsequent mini-batch (successive mini batches sequentially traverse the training set). The size of each mini-batch is 32. We train models using a learning rate of 0.002 and clip gradients with threshold 10.

B A list of remained words on IMDB

SparseVD with multiplicative weights retained the following words on IMDB task (sorted by descending frequency in the whole corpus):

start, oov, and, to, is, br, in, it, this, was, film, t, you, not, have, It, just, good, very, would, story, if, only, see, even, no, were, my, much, well, bad, will, great, first, most, make, also, could, too, any, then, seen, plot, acting, life, over, off, did, love, best, better, i, If, still, man, something, m, re, thing, years, old, makes, director, nothing, seems, pretty, enough, own, original, world, series, young, us, right, always, isn, least, interesting, bit, both, script, minutes, making, 2, performance, might, far, anything, guy, She, am, away, woman, fun, played, worst, trying, looks, especially, book, DVD, reason, money, actor, shows, job, 1, someone, true, wife, beautiful, left, idea, half, excellent, 3, nice, fan, let, rest, poor, low, try, classic, production, boring, wrong, enjoy, mean, No, instead, awful, stupid, remember, wonderful, often, become, terrible, others, dialogue, perfect, liked, supposed, entertaining, waste, His, problem, Then, worse, definitely, 4, seemed, lives, example, care, loved,

Why, tries, guess, genre, history, enjoyed, heart, amazing, starts, town, favorite, car, today, decent, brilliant, horrible, slow, kill, attempt, lack, interest, strong, chance, wouldn, sometimes, except, looked, crap, highly, wonder, annoying, Oh, simple, reality, gore, ridiculous, hilarious, talking, female, episodes, body, saying, running, save, disappointed, 7, 8, OK, word, thriller, Jack, silly, cheap, Oscar, predictable, enjoyable, moving, Unfortunately, surprised, release, effort, 9, none, dull, bunch, comments, realistic, fantastic, weak, atmosphere, apparently, premise, greatest, believable, lame, poorly, NOT, superb, badly, mess, perfectly, unique, joke, fails, masterpiece, sorry, nudity, flat, Good, dumb, Great, D, wasted, unless, bored, Tony, language, incredible, pointless, avoid, trash, failed, fake, Very, Stewart, awesome, garbage, pathetic, genius, glad, neither, laughable, beautifully, excuse, disappointing, disappointment, outstanding, stunning, noir, lacks, gem, F, redeeming, thin, absurd, Jesus, blame, rubbish, unfunny, Avoid, irritating, dreadful, skip, racist, Highly, MST3K

An interesting observation is that small and high ratings of films are kept while the middle values (5, 6) of the ratings are dropped.