**知能プログラミング演習I 第7回レポート**

2024年 7月 17 日 学籍番号 34714037 氏名 加藤薫

１．

(a)

,

(c)

(d)

(e),

(f),

(g)dim=0の場合は最初の次元が異なり行と列の位置が同じ要素についてソフトマックスが計算されるのですべての要素で0.5となる．dim=1の場合は各列に対してソフトマックスが計算され，各列の合計が１になるのですべての要素で0.25となる．dim=2の場合は各行に対してソフトマックスが計算され，各行の合計が１になるのですべての要素で0.33となる．

２．

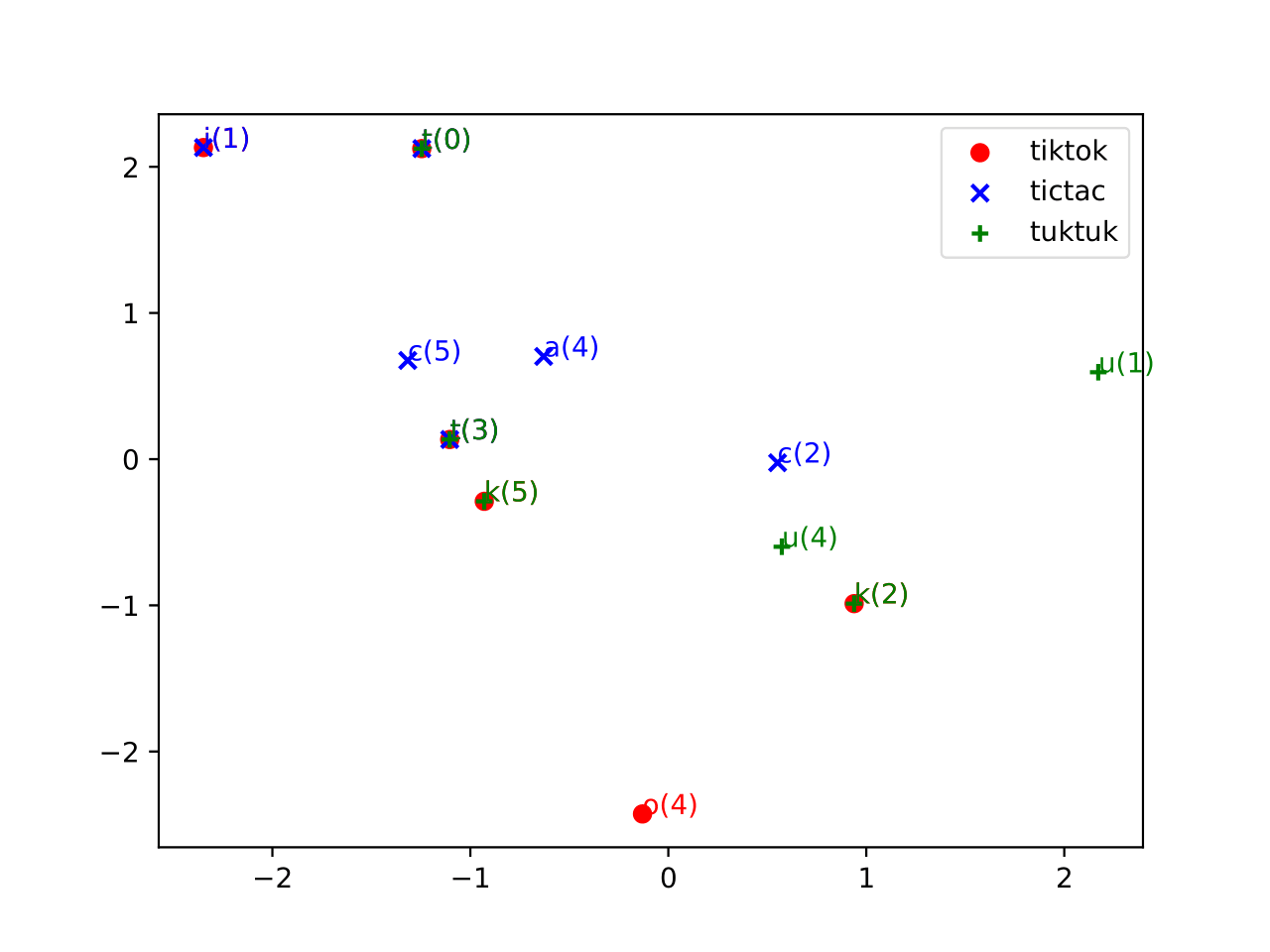
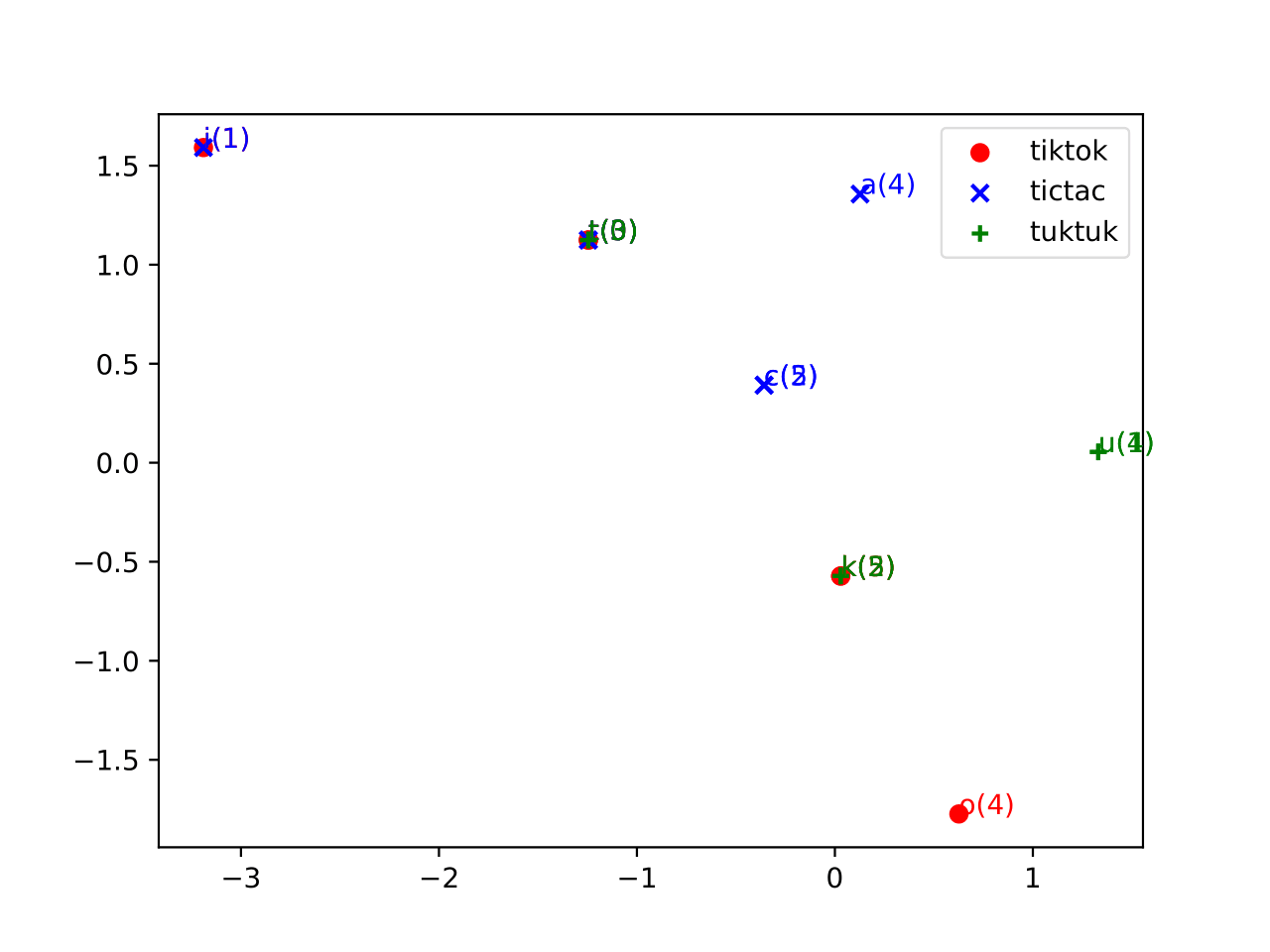
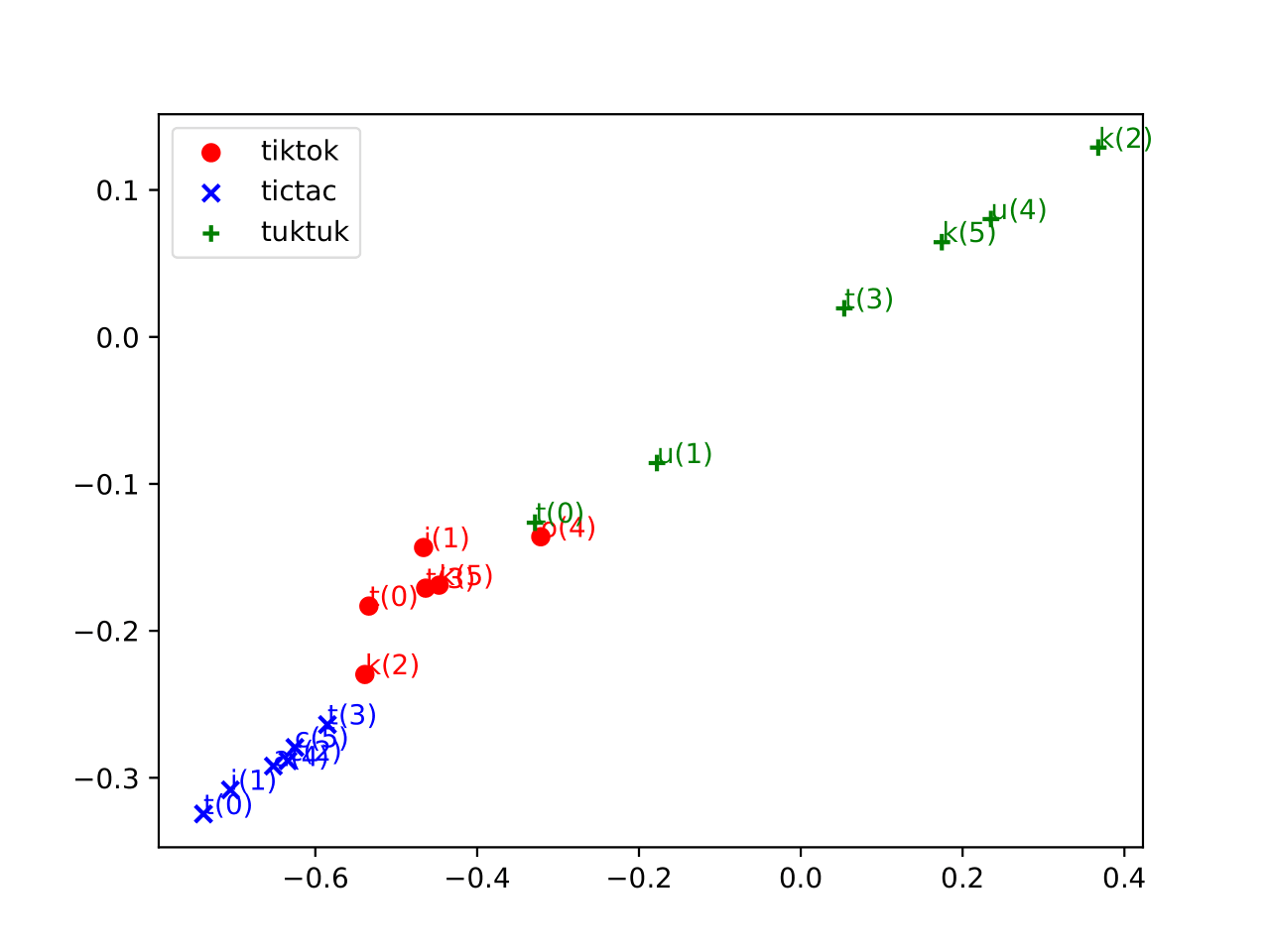
(a)文字列の各文字を，その文字に対応する数値で表現している．配列のサイズは3×６であり，1次元目が文字列の個数，2次元目が文字列の長さに対応している．

(b) token embeddingは文字の種類char sizeと埋め込み空間の次元数dim embedを引数から受け取り，文字を埋め込み表現の数値ベクトルに変換する．xのサイズは3×6×２であり，1次元目が文字列の個数，2次元目が文字列の長さ，３次元目が埋め込み空間の次元数である．

(c) embed-1.pdf の図は各文字の埋め込み表現の数値ベクトルを2次元空間にプロットしたものである．文字の種類が同じであれば，文字列や位置に関わらず，重複して同じ位置にプロットされる．

(d) embed-2.pdf の図は各文字の埋め込み表現の数値ベクトルをPositional Encoding の後で2次元空間にプロットしたものである．文字の種類と，前から何番目の文字かという位置も同じであれば重複して同じ位置にプロットされるが，位置が異なると異なる位置にプロットされる．

(e) embed-3.pdf の図は各文字の埋め込み表現の数値ベクトルをPositional Encoding の後Self-attention の順伝播計算をして2次元空間にプロットしたものである．文字の種類が同じであっても重複して同じ位置にプロットされることはなく，同じ文字列に含まれている文字同士で距離が小さく，文字列でグループ分けできる．t，i，tの3文字が同じtiktokとtictacグループの距離が小さいなど，似たグループの距離が近くなる傾向がある．直線状にプロットされた．

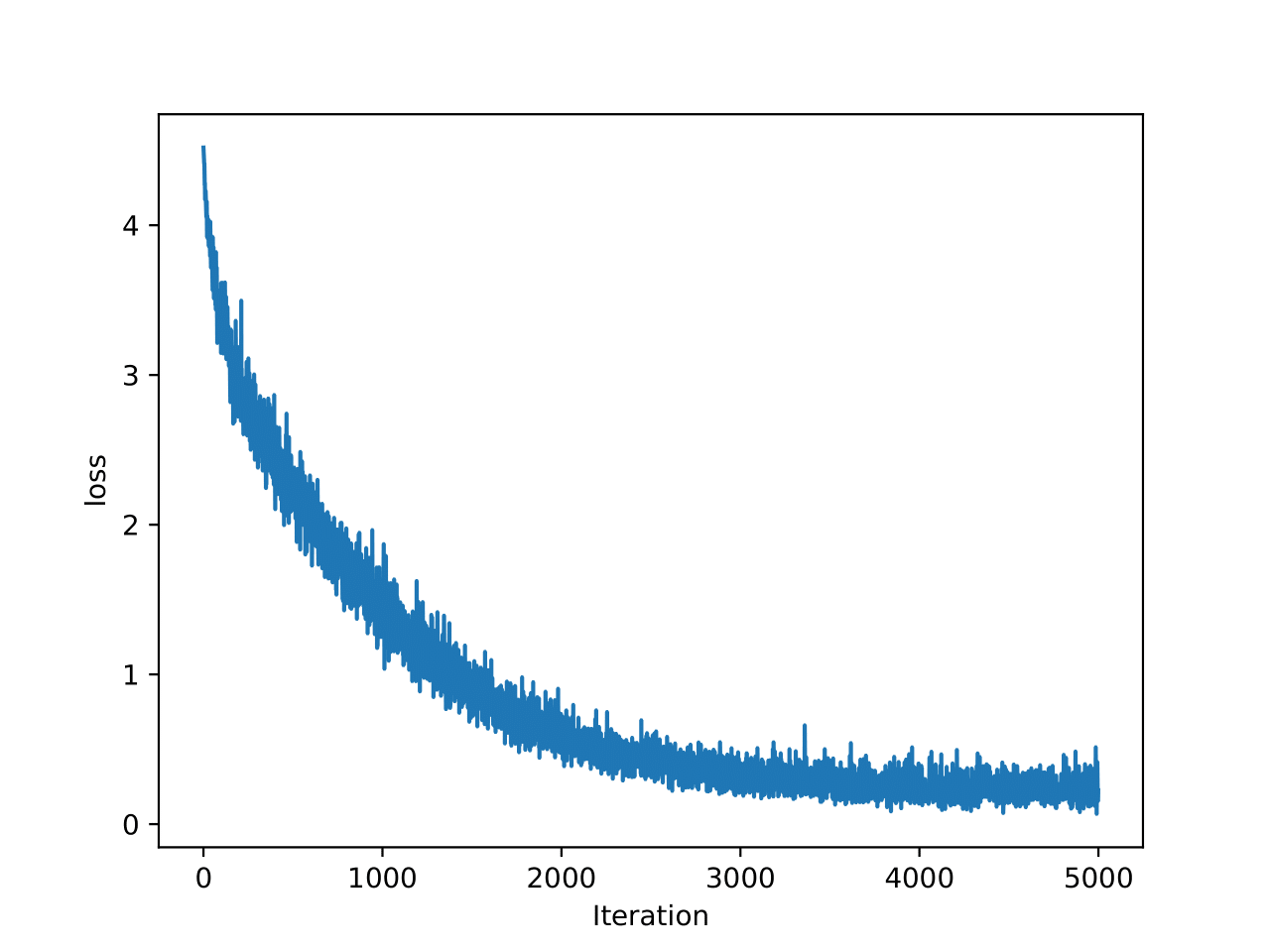
embed-1.pdf

embed-1.pdf embed-2.pdf embed3.pdf

３．

（a）loss fn(logits, y)は訓練データ学習時のクロスエントロピー損失，つまりトランスフォーマーの出力結果と入力文字列の最後の文字から一文字だけずれたものの誤差を計算している．

（b）損失の推移loss.pdfの図と，学習後に実行して生成された文章の例



入力されたプロンプト: ももからうまれた

NeuralNet出力(１回目) > ももからうまれたおとこのこをひろいあいないじょうにつるして、おばあさんをどおいばあさんがたくれますと、こうへかぬする

NeuralNet出力(２回目) > ももからうまれたおとこのこをひひめってこたぬまはたでするんいのぬきに、おじいさんはほしのかまいました。こかしです。で

入力されたプロンプト: せかいいちつよい

NeuralNet出力(１回目) > せかいいちつよいかぜさん。ちゅーこをおよめにもらってくれませんか？」「そりゃうれしいが、つくさんをもってもおいしいの

NeuralNet出力(２回目) > せかいいちつよいかべさん。ちゅーこをおよめにもらってくれませんか？」「そりゃうれしいが、かぜはわしよりつよいぞ。わし

入力されたプロンプト: おとこがこどもに

NeuralNet出力(１回目) > おとこがこどもにあぶをむすんだわかものは、みからいもつよいばんみかたのをやすってきまました。ちゅーい、よにもめものが