

emb_note

January 23, 2024

```
[ ]: import pandas as pd
import numpy as np
```

```
[ ]: df = pd.read_pickle("./sent_df.pkl")
df
```

```
[ ]:
      sent                                embedding    umap_0 \
0  This is it  [-0.02963203378021717, -0.0698486939072609, -0...  15.915230
1      This  [-0.023993512615561485, 0.024364134296774864, ...  15.966436
2      is    [0.019494930282235146, -0.0025707106105983257,...  15.423137
3      it    [0.0017801866633817554, 0.05635393410921097, -...  16.460556
4      This  [-0.023993512615561485, 0.024364134296774864, ...  15.422277
5  This is  [-0.00482132937759161, -0.053047049790620804, ...  14.841746
6  This is it  [-0.02963203378021717, -0.0698486939072609, -0...  16.645634

      umap_1
0  10.111320
1  11.961027
2  10.968864
3  12.551167
4  12.703674
5  11.719730
6  10.766572
```

```
[ ]: np.array(df["embedding"][1:4].to_list())[:,0].sum()
```

```
[ ]: lm_df = pd.read_pickle("./sent_llama.pkl")
lm_df
```

```
[ ]:
      sent                                embedding    umap_0 \
0  This is it  [2.804270029067993, -3.0918750762939453, 1.886...  14.612863
1      This  [1.4891676902770996, 1.6863242387771606, 1.776...  16.436907
2      is    [-0.0609406940639019, -1.0181427001953125, 0.9...  16.470743
3      it    [-0.3512643277645111, 2.036311149597168, 1.649...  17.161087
4      This  [1.4891676902770996, 1.6863242387771606, 1.776...  15.489168
5  This is  [1.3410488367080688, -2.2461018562316895, 1.57...  15.173997
6  This is it  [2.804270029067993, -3.0918750762939453, 1.886...  15.432034
```

```

    umap_1
0  4.923532
1  5.855392
2  6.890014
3  6.666708
4  6.738914
5  5.716593
6  4.791289

```

```
[ ]: lm_df = pd.read_pickle("./sent_mst.pkl")
lm_df
```

```
[ ]:
      sent                                embedding  umap_0 \
0  This is it  [0.5282784700393677, -5.262070178985596, 1.427...  7.825799
1      This  [-0.40162912011146545, 2.2741453647613525, 0.6...  7.163516
2      is  [-2.6447949409484863, 1.8061383962631226, -1.3...  9.366054
3      it  [-4.764896869659424, -3.2947678565979004, 6.07...  9.655240
4      This  [-0.40162912011146545, 2.2741453647613525, 0.6...  7.376906
5  This is  [-1.2165313959121704, 4.944725036621094, 6.539...  8.213544
6  This is it  [0.5282784700393677, -5.262070178985596, 1.427...  8.692219

```

```

    umap_1
0  19.304983
1  18.006935
2  17.983381
3  18.783037
4  17.118874
5  18.270605
6  19.237478

```

```
[ ]:
```