code-embedding

July 15, 2025

```
[1]: from transformers import AutoTokenizer, AutoModel
     import torch
```

```
[3]: model_name = "huggingface/CodeBERTa-small-v1"
     tokenizer = AutoTokenizer.from_pretrained(model_name)
     model = AutoModel.from_pretrained(model_name)
```

```
/usr/local/lib/python3.11/dist-packages/huggingface_hub/utils/_auth.py:94:
UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab
(https://huggingface.co/settings/tokens), set it as secret in your Google Colab
and restart your session.
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access
public models or datasets.
  warnings.warn(
```

tokenizer_config.json: 0%| | 0.00/19.0 [00:00<?, ?B/s]

config.json: 0%| | 0.00/480 [00:00<?, ?B/s]

vocab.json: 0.00B [00:00, ?B/s]

merges.txt: 0.00B [00:00, ?B/s]

pytorch_model.bin: 0%| | 0.00/336M [00:00<?, ?B/s]

```
[8]: user_input = input("Enter a description or code snippet: ")
```

Enter a description or code snippet: sort dictionary in python

```
[9]: inputs = tokenizer(user_input, return_tensors="pt")
```

```
[10]: with torch.no_grad():
          outputs = model(**inputs)
          # Get the last hidden state
          embeddings = outputs.last_hidden_state.mean(dim=1).squeeze()
```

```
[11]: print(f"\n Input: {user_input}")
      print(f" Embedding (vector of size {embeddings.shape[0]}):")
      print(embeddings)
```

```
 Input: sort dictionary in python
 Embedding (vector of size 768):
tensor([-3.9466e-01,  3.6335e-01, -8.6741e-01,  2.8478e-01, -1.9209e-02,
         3.1392e-01,  5.6974e-01, -1.0654e+00, -1.2790e+00,  4.3474e-01,
         1.6026e+00, -9.0060e-01,  1.1317e+00, -1.5454e+00, -1.1173e+00,
        -9.9257e-01, -8.2647e-01,  1.2201e-01,  9.7324e-01, -4.4252e-01,
        -1.0289e+00,  1.4818e-01, -4.1146e-01,  1.4719e+00,  4.7343e-01,
        -5.0719e-01,  4.7786e-01,  3.2028e-01, -6.8621e-01, -1.9049e-01,
         8.6703e-01, -6.4549e-02,  1.0746e-01, -1.0194e+00,  8.2436e-01,
        -7.3307e-01,  3.5306e-01, -1.8640e-01,  6.1714e-01,  3.8319e-01,
        -1.5574e+00, -3.5061e-01,  9.0753e-01,  7.8903e-01, -3.7470e-01,
        -4.5818e-01,  2.2368e-01, -5.5922e-01, -5.8966e-02, -2.0671e-01,
         4.4980e-01, -6.7169e-01, -1.0539e+00, -2.0786e-01,  1.1640e+00,
         5.7450e-01,  5.6518e-01,  1.6835e-01,  4.9969e-01,  1.6888e+00,
         4.8128e-01, -1.8417e-01,  8.0044e-01,  6.6567e-01,  7.2740e-01,
         2.2254e+00,  4.7164e-02,  8.3773e-01,  5.3767e-01,  1.0450e-01,
         9.0650e-01, -1.0208e+00, -1.6950e-01, -1.6491e+00, -4.1452e-01,
         5.2064e-01,  2.0346e-01,  5.2197e-03, -3.1903e-01, -1.4833e+00,
        -7.5583e-01,  1.3802e-01,  1.3846e+00, -2.7681e-01, -8.0368e-01,
         5.0220e-01, -9.6343e-01, -7.4861e-01, -2.7681e-01,  2.6921e-01,
        -2.2263e-03, -1.2095e+00,  2.5607e-02, -2.4832e-01, -1.4428e-01,
        -1.0385e+00, -6.4950e-01, -8.2672e-01, -5.5683e-01,  8.3759e-03,
         6.1570e-01,  8.5335e-03,  7.5600e-01, -4.3919e-01, -6.6216e-01,
         1.1537e-01, -6.6438e-01,  2.2828e-01,  7.9391e-01,  5.1069e-01,
        -1.1437e+00,  1.7905e+00, -6.4498e-01, -7.7108e-01, -4.2902e-03,
        -5.9760e-01,  1.0069e-01,  4.0338e-01,  5.4362e-01,  3.3229e-01,
         1.3182e+00,  9.1007e-02, -5.3462e-01,  2.0684e+00,  2.2826e-01,
         4.2456e-01, -3.4519e-01, -9.3099e-01, -1.2464e+00, -1.6536e-01,
        -9.2179e-01,  9.8243e-01,  7.3609e-01,  1.1417e+00,  2.1472e-01,
         1.5773e+00, -5.0548e-01,  2.4271e-01, -6.1112e-01,  3.0681e-01,
        -5.0964e-02, -1.9903e+00,  4.6539e-01,  6.1441e-01,  1.6510e-01,
         2.4289e-01, -7.1153e-01,  8.2525e-01, -6.0442e-01,  8.6325e-01,
         8.1648e-02,  4.9499e-01, -4.4694e-01, -4.5629e-01,  1.4572e+00,
         4.1698e-01,  9.7344e-02,  4.1735e-01, -8.6787e-01,  3.1454e-01,
         1.8011e+00,  5.5043e-01, -1.2030e+00, -3.4548e-03, -5.9522e-02,
        -1.0373e+00, -9.5468e-01,  2.4547e-01,  3.7625e-01,  3.3307e-01,
         4.7441e-01,  5.3706e-01, -7.3987e-01, -2.6119e-01,  5.4067e-02,
         6.8697e-01,  1.3366e+00,  9.4656e-02,  8.2307e-02,  1.5038e+00,
        -2.0080e-01, -3.2062e-01,  7.6866e-01,  9.1275e-02,  7.6121e-01,
        -5.7151e-01,  6.2829e-01, -8.9210e-01,  1.2283e+00, -4.7967e-01,
         2.8101e-01, -1.0692e+00, -1.7105e-01, -7.3534e-02, -1.0592e+00,
         1.4033e-01,  6.9687e-01,  9.7709e-01,  5.4242e-01,  3.6697e-01,
```

```
-1.8359e-01,  1.4754e+00,  1.9696e+00,  4.3815e-01, -3.2235e-02,
 5.9711e-02, -1.5435e-01,  3.8712e-01, -4.0549e-01,  1.9449e-01,
 5.4256e-02, -1.3736e-01, -4.5622e-01,  5.0428e-01,  1.8111e+00,
-5.5569e-01, -3.8344e-01, -4.8246e-01, -8.7673e-01, -6.0934e-01,
 4.9536e-01, -5.3753e-01, -3.3901e-01,  8.1206e-02, -2.1566e-01,
-8.0439e-01,  5.8207e-01,  4.2404e-01, -2.1615e-01,  4.1644e-01,
-1.1772e+00,  3.4418e-01, -2.4365e-01,  1.7492e-01, -5.4352e-01,
 2.3576e-01,  1.0808e-01, -3.9403e-01,  2.9290e-01, -1.6484e+00,
-1.8416e+00, -4.2567e-01, -1.9047e-01,  5.0544e-01, -4.6132e-01,
 6.3622e-01, -4.9497e-01,  9.9988e-01, -3.7348e-02,  3.9907e-02,
-1.2805e+00, -2.8996e-01, -7.6912e-01, -3.1081e-01,  4.4994e-01,
-8.5815e-01, -1.2347e+00, -6.1843e-01,  3.1678e-01,  2.8267e-01,
 3.4278e-01, -1.2160e+00, -1.3996e+00,  1.7570e+00,  1.5155e-01,
-7.9765e-01,  1.7658e+00, -1.2506e+00, -6.8644e-01, -5.4368e-01,
 1.1183e+00, -1.0175e+00, -3.4512e-01, -1.1876e+00, -1.3419e+00,
-2.0341e-01,  1.6815e-01, -6.2958e-01, -1.2874e+00,  1.4726e+00,
 8.0487e-01, -1.0996e-01, -1.0359e-01,  1.0914e+00,  6.8475e-01,
 9.5486e-01,  5.8191e-01,  3.9849e-01,  5.8347e-02,  3.0792e-01,
-1.8492e-02,  1.5210e-01,  4.7283e-01,  1.9465e-01, -7.6870e-01,
-8.3691e-02, -4.0002e-01, -7.3165e-02, -9.0742e-01,  1.0877e+00,
 1.0916e+00, -1.6852e+00,  1.7015e-01, -7.0005e-01, -5.8583e-01,
-6.9784e-01,  7.1735e-03,  7.1111e-01,  3.1732e-01, -1.1939e-01,
-6.9848e-01,  6.5678e-01, -6.2836e-01,  3.1192e-01,  4.2905e-01,
-1.5862e+00,  6.6990e-01, -1.0189e+00, -1.0171e+00, -4.4217e-01,
-3.8576e-01,  6.0507e-02, -3.6012e-01,  1.1897e+00, -7.7542e-02,
 8.9362e-01,  3.7781e-01, -6.6249e-01, -1.1341e+00, -3.8466e-01,
 1.1964e+00,  2.2420e-01, -2.6287e+00, -2.7936e-01,  8.5628e-01,
-2.1809e-01,  8.8043e-01, -7.6416e-01,  5.5032e-01,  3.6902e-01,
 1.0430e+00, -1.5165e-01,  3.1623e-02, -1.2060e+00,  2.9662e-01,
-5.6501e-01, -4.8290e-01,  1.3952e+00,  1.8780e+00,  5.6753e-01,
-3.1503e-01, -2.7606e-01,  7.7368e-01,  1.4221e+00, -8.9550e-01,
-3.2714e-01, -1.0837e+00,  1.6431e+00,  1.8470e+00, -7.6060e-01,
-4.7113e-01, -4.1350e-01, -8.6469e-01, -1.1767e+00, -2.5136e-01,
-1.8446e+00, -8.3531e-01, -7.4407e-01,  2.1719e-01, -4.0324e-01,
 1.0463e-02,  2.7335e-01, -1.4794e-01,  1.3882e+00,  5.2093e-02,
-6.1117e-01,  1.0295e+00,  5.0619e-01,  1.0609e+00,  1.7449e-01,
-4.6366e-01,  4.6918e-03, -5.9782e-01,  1.4693e+00,  8.1530e-01,
 6.4223e-01,  1.1400e+00,  9.3435e-01, -1.7464e+00, -1.9523e-01,
-1.8438e-01,  8.8885e-01,  3.9995e-02,  1.3509e+00, -4.3118e-01,
-4.1160e-02,  1.8058e+00, -6.8455e-02,  2.7726e-01, -3.5626e-01,
 6.1691e-01, -9.0448e-01, -1.1325e-01,  6.3989e-01,  5.6167e-01,
-5.2537e-01, -1.0787e+00,  6.2494e-02,  4.2996e-02, -1.3825e+00,
 6.3659e-01,  4.0412e-01, -3.1214e-01, -1.4068e+00,  1.0365e-01,
 6.3250e-01,  5.1903e-01, -8.1086e-01, -8.6198e-01, -6.4649e-01,
 1.2571e+00,  1.3052e-01, -3.0795e-01,  1.8117e+00,  1.6622e+00,
 8.9887e-01, -8.0652e-01,  1.9701e-01,  9.7951e-01, -2.4013e-01,
-2.3349e-02,  3.6430e-01, -9.1191e-01, -3.4236e-01, -7.5788e-01,
-4.8252e-01, -1.4521e+00, -8.9078e-01,  6.4862e-01,  5.4360e-02,
```

```
-5.5253e-01,  9.1309e-03,  1.1272e+00,  1.2504e-01, -6.3468e-01,
 1.2206e+00,  7.6323e-01,  6.9973e-02,  8.0182e-01, -3.6480e-01,
 3.7936e-01,  2.5246e-01, -4.3393e-01,  8.3660e-01, -1.7287e+00,
-1.1519e+00, -3.9188e-01,  6.6286e-01, -1.0397e+00, -4.4948e-01,
-7.4875e-01, -7.9737e-01, -9.0874e-01,  1.2162e-01, -8.7131e-01,
 7.5239e-01,  9.2484e-01,  1.9512e-01, -2.2891e+00,  1.7615e-02,
-2.8520e-02,  1.0298e-02,  6.7428e-01, -9.9878e-02,  6.7796e-01,
-7.2733e-01, -1.3563e+00, -6.8626e-01, -5.9259e-01,  1.0597e-01,
 1.3785e+00, -1.8837e-01,  2.3108e-02, -8.7430e-01,  7.8322e-01,
 8.7348e-01,  8.0510e-01, -2.4606e-01,  7.7856e-01,  1.3518e+00,
-1.1211e+00, -7.6478e-01,  1.7686e+00, -6.3837e-01, -3.8158e-01,
 1.0113e+00,  1.0740e-01,  5.1465e-01,  2.5606e-01, -1.7131e+00,
 9.5230e-01,  2.0904e+00,  4.0811e-01,  2.6280e-01,  6.1729e-01,
-6.1043e-01,  6.8550e-01,  1.3759e+00,  9.1966e-01,  7.9977e-01,
-9.1187e-01,  6.8181e-03, -8.2189e-01,  9.9764e-02,  1.5296e+00,
 1.6554e-01, -4.8372e-01,  7.3923e-01, -1.3240e+00, -8.3051e-01,
-5.3933e-01, -7.9755e-01,  3.2184e-01, -8.1177e-01, -3.4583e-01,
-9.9160e-02, -1.3007e-01, -9.3751e-01,  5.3582e-01,  2.1026e-01,
-9.9297e-01, -2.1177e-01,  6.0491e-01,  4.8730e-01,  1.1919e+00,
-1.0238e+00,  3.9433e-01,  8.8262e-01, -1.7458e-01, -8.5806e-02,
-4.1216e-02, -3.2888e-01,  1.3495e+00, -3.5903e-01, -1.5912e-01,
-7.8823e-01, -6.8049e-01,  1.0708e+00,  2.4436e-01, -3.9852e-01,
-4.8128e-02, -1.7591e+00, -7.5647e-01, -1.0247e-01, -7.3030e-01,
-9.6419e-01,  1.3916e+00,  5.4080e-01, -4.6260e-01,  6.5846e-02,
 4.2399e-01, -6.4573e-01, -3.7037e-01,  5.4456e-01,  4.8065e-02,
-4.3778e-02, -3.1228e-01, -6.3123e-02, -2.7684e-01, -4.9404e-01,
-1.3368e+00,  1.0594e+00, -2.1732e-01, -1.8054e-01,  4.0188e-01,
 1.8838e-01, -9.0442e-01,  4.6268e-01, -5.1246e-01,  4.2868e-01,
 1.5360e-01, -1.0076e+00, -7.1081e-02,  6.4104e-01,  9.0886e-01,
-2.8360e-01, -4.0044e-01, -5.9759e-01, -2.8583e-01, -5.7885e-01,
-6.9464e-02, -5.3071e-01, -7.0163e-02, -1.7513e-01,  4.4745e-01,
 2.0284e-01, -5.8474e-01,  1.9396e-01, -4.5077e-01, -3.8819e-01,
-6.8351e-01,  4.3312e-03,  1.4036e-01,  1.7829e+00,  4.6848e-01,
 7.7336e-01,  3.1831e-01, -3.5402e-02, -1.4926e+00,  1.8893e-01,
 8.7497e-01,  3.5357e-01,  4.8834e-01, -6.0301e-02, -1.6783e-01,
-1.2551e-01, -6.7605e-01,  4.1875e-01, -8.2666e-01, -5.0009e-01,
 5.2265e-01,  7.0102e-01,  4.3874e-01, -2.7241e-01,  5.3859e-02,
-9.8074e-01, -1.6869e-01,  2.2484e-01, -1.2755e-01, -3.5726e-01,
-4.7851e-01, -7.4401e-01,  5.6647e-01,  2.4131e-01,  1.4771e-01,
-8.9022e-02,  5.8515e-01,  5.1442e-01, -9.8485e-03,  1.4485e-02,
-2.4877e-01, -7.4972e-01,  1.0610e+00,  6.2169e-01,  1.5768e-01,
 7.1900e-01,  8.0905e-02,  3.6740e-01,  5.6215e-02, -5.4610e-01,
-8.3193e-01, -5.2741e-01, -4.0545e-01,  7.3850e-01, -1.6472e+00,
 7.6582e-02, -1.6550e-01,  5.1118e-01,  6.0523e-01,  1.4081e-01,
-9.8484e-01,  1.6152e+00, -1.3452e+00, -1.8022e+00, -1.4820e+00,
 1.7633e-01,  4.5024e-01, -1.0988e+00, -3.3319e-01, -5.9800e-01,
-5.7340e-01,  1.1231e+00,  7.3628e-01,  1.7446e-01,  2.1999e-01,
-3.7923e-01, -1.1968e+00, -2.8678e-01,  2.7111e-01,  5.4801e-01,
```

```
-1.8777e-01, -5.0493e-01, -4.1693e-01,  1.5824e+00,  7.9618e-01,
 1.8968e-01,  1.1258e+00,  7.9184e-01,  3.4770e-01, -6.7755e-01,
-6.1851e-01,  1.1058e+00,  1.6672e-01, -1.6645e+00,  3.3301e-01,
 7.6593e-01,  9.8830e-01, -3.1667e-01,  5.3534e-01, -5.6463e-01,
-1.5383e-01,  1.7672e-01, -1.6494e-02, -1.0236e+00,  5.6947e-01,
 2.0634e-01, -1.7643e+00, -5.2710e-02, -3.4725e-01,  6.3621e-01,
 3.7968e-01,  1.0220e+00,  1.5025e+00, -8.6825e-01,  1.2750e+00,
 3.8087e-01,  3.1238e-01,  1.1599e+00,  7.0763e-01, -4.2718e-01,
 2.5958e-01,  4.9259e-01, -6.7616e-01, -4.7650e-01,  2.1084e-03,
-3.9914e-01,  3.4470e-01, -3.3885e-01,  1.7333e+00,  4.7447e-01,
-7.8158e-02,  8.7411e-01, -3.0268e-01, -1.8335e+00,  1.1683e-03,
-8.6297e-01,  6.4281e-01,  6.2864e-01,  1.3163e+00, -5.2347e-01,
-3.7585e-01, -2.1490e-01,  7.0624e-01, -1.5150e+00, -5.6380e-01,
-8.7637e-02, -1.1085e-02, -2.4764e-01, -2.0119e-01, -4.3114e-01,
 2.7759e-01,  1.2968e+00, -2.6161e-01,  5.4572e-01,  6.1150e-01,
 3.7596e-01,  7.0634e-01, -2.4538e-01, -3.5238e-01, -7.6065e-01,
-2.4694e-01, -7.4111e-01, -6.1794e-03,  4.1471e-01,  1.6654e-02,
-8.1195e-01,  3.8268e-01,  5.6363e-01]])
```