

```
1 E:\Desktop\000\000\venv\Scripts\python.exe "D:\Bristol course\Data Science Dissertation\ExpBERT lab and thesis\BERT-IE-main\BERT-IE-main\diversity_sampling_whole.py"
2 diversity_sampling_openai
3 Downloading and preparing dataset csv/default to C:/Users/admin/.cache/huggingface/datasets/csv/default-2067039c60457556/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1...
4 Downloading data files: 100%|██████████| 1/1 [00:00<00:00, 2015.52it/s]
5 Extracting data files: 100%|██████████| 1/1 [00:00<00:00, 7.48it/s]
6 0%|          | 0/1 [00:00<?, ?it/s]Dataset csv downloaded and prepared to C:/Users/admin/.cache/huggingface/datasets/csv/default-2067039c60457556/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1. Subsequent calls will reuse this data.
7 100%|██████████| 1/1 [00:00<00:00, 19.95it/s]
8 11241
9 torch.Size([11241, 27])
10 Downloading and preparing dataset csv/default to C:/Users/admin/.cache/huggingface/datasets/csv/default-cb4bfe0feb86dc8e/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1...
11 Downloading data files: 100%|██████████| 1/1 [00:00<?, ?it/s]
12 Extracting data files: 100%|██████████| 1/1 [00:00<00:00, 99.03it/s]
13 0%|          | 0/1 [00:00<?, ?it/s]Dataset csv downloaded and prepared to C:/Users/admin/.cache/huggingface/datasets/csv/default-cb4bfe0feb86dc8e/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1. Subsequent calls will reuse this data.
14 100%|██████████| 1/1 [00:00<00:00, 41.81it/s]
15 Downloading data files: 100%|██████████| 1/1 [00:00<?, ?it/s]
```

```
16 Extracting data files: 100%|██████████| 1/1 [00:00<00:00, 125.33it/s]
17 Generating train split: 0 examples [00:00, ? examples /s]Downloading and preparing dataset csv/default to C:/Users/admin/.cache/huggingface/datasets/csv/default-5fcc223563fb6e77/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1...
18 100%|██████████| 1/1 [00:00<00:00, 59.02it/s]
19 Dataset csv downloaded and prepared to C:/Users/admin/.cache/huggingface/datasets/csv/default-5fcc223563fb6e77/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1. Subsequent calls will reuse this data.
20 noexp_df.shape[0] last 11241
21 unexplained dataset length is:
22 11241
23 embedding begin
24 trian and val dataset length: 28350
25 diversity_sampling_openai
26 diversity_sampling_openai
27 diversity_sampling_openai
28 diversity_sampling_openai
29 diversity_sampling_openai
30 diversity_sampling_openai
31 diversity_sampling_openai
32 diversity_sampling_openai
33 315.0
34 ./embeddings/NEW_bertie_embeddings_textattack/bert-base-uncased-MNLI_subset_1.pt
35 Downloading and preparing dataset csv/default to C:/Users/admin/.cache/huggingface/datasets/csv/default-494a068ef8ae599d/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1...
36 Downloading data files: 100%|██████████| 1/1 [00:00<00:00, 2252.58it/s]
37 Extracting data files: 100%|██████████| 1/1 [00:00<00:00, 394.91it/s]
38 100%|██████████| 1/1 [00:00<00:00, 58.45it/s]
39 Dataset csv downloaded and prepared to C:/Users/admin
```

```
39 ./cache/huggingface/datasets/csv/default-
494a068ef8ae599d/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc853
bcc7ec513e1. Subsequent calls will reuse this data.
40 test len 14445
41 Downloading data files: 100%|██████████| 1/1 [00:00<
00:00, 1003.66it/s]
42 Extracting data files:  0%|          | 0/1 [00:00
<?, ?it/s]Downloading and preparing dataset csv/
default to C:/Users/admin/.cache/huggingface/datasets
/csv/default-050e89251c29af62/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc853
bcc7ec513e1...
43 Extracting data files: 100%|██████████| 1/1 [00:00<00
:00, 47.22it/s]
44 0%|          | 0/1 [00:00<?, ?it/s]Dataset csv
downloaded and prepared to C:/Users/admin/.cache/
huggingface/datasets/csv/default-050e89251c29af62/0.0
.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc853
bcc7ec513e1. Subsequent calls will reuse this data.
45 100%|██████████| 1/1 [00:00<00:00, 37.16it/s]
46 preprocess_samples begin ...
47 uncertainty shape: 1530
48 ./test_embeddings/NEW_bertie_embeddings_textattack/
bert-base-uncased-MNLI_subset_1.pt
49 classify begin
50 3211
51 0%|          | 0/20 [00:00<?, ?it/s]
    diversity_sampling_openai
52 diversity_sampling_openai
53 320it [00:08, 96.93it/s]diversity_sampling_openai
54 diversity_sampling_openai
55 diversity_sampling_openai
56 diversity_sampling_openai
57 647it [00:21, 50.72it/s]diversity_sampling_openai
58 diversity_sampling_openai
59 diversity_sampling_openai
60 diversity_sampling_openai
61 1054it [00:34, 63.69it/s]diversity_sampling_openai
62 diversity_sampling_openai
```

```
63 diversity_sampling_openai
64 1066it [00:50, 10.56it/s]diversity_sampling_openai
65 1420it [00:51, 70.02it/s]diversity_sampling_openai
66 diversity_sampling_openai
67 diversity_sampling_openai
68 1421it [01:06, 9.46it/s]diversity_sampling_openai
69 1739it [01:07, 53.02it/s]diversity_sampling_openai
70 diversity_sampling_openai
71 diversity_sampling_openai
72 diversity_sampling_openai
73 2095it [01:20, 50.92it/s]diversity_sampling_openai
74 diversity_sampling_openai
75 diversity_sampling_openai
76 diversity_sampling_openai
77 2461it [01:33, 51.27it/s]diversity_sampling_openai
78 diversity_sampling_openai
79 diversity_sampling_openai
80 2486it [01:45, 14.17it/s]diversity_sampling_openai
81 2747it [01:46, 46.27it/s]diversity_sampling_openai
82 diversity_sampling_openai
83 diversity_sampling_openai
84 diversity_sampling_openai
85 3183it [01:59, 68.22it/s]diversity_sampling_openai
86 diversity_sampling_openai
87 diversity_sampling_openai
88 diversity_sampling_openai
89 3538it [02:12, 52.90it/s]diversity_sampling_openai
90 diversity_sampling_openai
91 diversity_sampling_openai
92 diversity_sampling_openai
93 3881it [02:24, 49.56it/s]diversity_sampling_openai
94 diversity_sampling_openai
95 diversity_sampling_openai
96 3906it [02:37, 13.77it/s]diversity_sampling_openai
97 4242it [02:37, 63.90it/s]diversity_sampling_openai
98 diversity_sampling_openai
99 diversity_sampling_openai
100 diversity_sampling_openai
101 4590it [02:50, 51.60it/s]diversity_sampling_openai
102 diversity_sampling_openai
103 diversity_sampling_openai
```

```
104 diversity_sampling_openai
105 4954it [03:03, 51.28it/s]diversity_sampling_openai
106 diversity_sampling_openai
107 diversity_sampling_openai
108 4971it [03:16, 13.46it/s]diversity_sampling_openai
109 5299it [03:16, 61.65it/s]diversity_sampling_openai
110 diversity_sampling_openai
111 diversity_sampling_openai
112 diversity_sampling_openai
113 5648it [03:29, 50.32it/s]diversity_sampling_openai
114 diversity_sampling_openai
115 diversity_sampling_openai
116 diversity_sampling_openai
117 5993it [03:42, 48.45it/s]diversity_sampling_openai
118 diversity_sampling_openai
119 diversity_sampling_openai
120 diversity_sampling_openai
121 6357it [03:55, 49.69it/s]diversity_sampling_openai
122 diversity_sampling_openai
123 diversity_sampling_openai
124 diversity_sampling_openai
125 6704it [04:09, 47.91it/s]diversity_sampling_openai
126 diversity_sampling_openai
127 diversity_sampling_openai
128 diversity_sampling_openai
129 7045it [04:22, 47.02it/s]diversity_sampling_openai
130 diversity_sampling_openai
131 epoch number 19
132 diversity_sampling_openai
133 diversity_sampling_openai
134 7100it [04:35, 47.02it/s]diversity_sampling_openai
135 diversity_sampling_openai
136 global t is : 0
137 test epoch results {'accuracy': 0.5392156862745098
, 'f1_weighted': 0.4939554215710343, 'f1_macro': 0.
37401773727144544, 'f1': array([0.80693069, 0
. , 0.26865672, 0.11594203, 0.51918736,
0.04347826, 0.5033557 , 0.58989229, 0.
51871658]), 'precision': array([0.77990431, 0
. , 0.5 , 0.57142857, 0.50218341,
0.66666667, 0.60483871, 0.47152318, 0.
```

```

139 54494382]), 'recall': array([0.83589744, 0
140 . , 0.18367347, 0.06451613, 0.53738318,
140 0.02247191, 0.43103448, 0.78761062, 0.
49489796])}
141 val epoch results {'accuracy': 0.5714285714285714, '
f1_weighted': 0.522547618545668, 'f1_macro': 0.
41390963004281167, 'f1': array([0.85365854, 0
. , 0.47058824, 0. , 0.52873563,
142 0.1 , 0.64615385, 0.61176471, 0.
51428571]), 'precision': array([0.83333333, 0
. , 0.66666667, 0. , 0.53488372,
143 1. , 0.72413793, 0.47852761, 0.6
]), 'recall': array([0.875 , 0.
0.36363636, 0. , 0.52272727,
144 0.05263158, 0.58333333, 0.84782609, 0.45
])}
145 7100it [04:37, 25.60it/s]
146 diversity len embedding 11241
147 MLP_1h(
148     (l1): Linear(in_features=27, out_features=100,
bias=True)
149     (l2): Linear(in_features=100, out_features=9, bias
=True)
150 )
151 [4828, 735, 355, 5575, 6248, 6722, 10424, 8715, 3216
, 2766, 8448, 4922, 6470, 5997, 1778, 4985, 6665,
9469, 8925, 3502, 5071, 6826, 1549, 3102, 1094, 6721
, 2240, 9846, 449, 5793, 4141, 5952, 2028, 10239,
6114, 8364, 10871, 5447, 4850, 4533, 2073, 4606,
8133, 6717, 2253, 222, 739, 1327, 6563, 362, 6684,
8485, 1607, 7607, 1575, 4273, 6455, 7559, 6535, 6856
, 2555, 3592, 5332, 7019, 233, 4912, 1084, 2478,
6274, 7427, 5910, 9475, 2097, 9266, 8824, 5816, 1611
, 9073, 9666, 9858, 3493, 2147, 1314, 5637, 2337,
2437, 3921, 3117, 10538, 9697, 6792, 1095, 6798,
5483, 7555, 514, 4708, 7032, 1350, 2977, 6730, 4979
, 4538, 1539, 3828, 4072, 1586, 7232, 4034, 5079,
1368, 10375, 5164, 8200, 10876, 1326, 9852, 3097,
3442, 3460, 11055, 2906, 2584, 3436, 10719, 2788,
3023, 1119, 9488, 1868, 53, 9747, 1304, 9314, 8948,
4563, 8414, 9400, 4086, 377, 10260, 3975, 10848,
```

151	4223, 3361, 3703, 9107, 1154, 7906, 4332, 8018, 2784 , 5153, 2901, 6203, 1147, 9166, 58, 7388, 5283, 778 , 7014, 2258, 8094, 5443, 4787, 2010, 4256, 422, 7373, 523, 4406, 841, 5127, 4037, 3587, 1121, 4175, 10709, 8623, 4264, 8662, 5117, 9769, 1794, 3169, 4868, 965, 6966, 2816, 3786, 6667, 2182, 332, 2321, 10789, 2091, 8862, 10381, 9193, 7994, 3875, 10735, 6454, 8656, 6912, 3073, 6951, 4166, 2402, 3340, 5177 , 3156, 9650, 5984, 10379, 2861, 7443, 7856, 5174, 6560, 10095, 976, 1963, 10587, 3204, 1150, 6406, 4627, 2373, 9374, 7204, 1183, 10458, 2641, 9413, 2254, 2283, 7111, 5018, 1628, 10435, 7505, 5368, 10621, 8707, 2597, 4119, 2431, 7013, 1035, 10970, 4904, 203, 1643, 7375, 9458, 3095, 9105, 5551, 2874 , 4703, 1122, 11191, 5137, 8700, 3247, 4859, 9516, 6765, 7625, 8406, 9598, 3907, 6311, 5415, 4347, 3878 , 8937, 3186, 4659, 1243, 3253, 318, 8890, 4757, 2556, 7759, 1464, 5478, 4643, 6608, 11193, 5815, 109 , 9649, 5955, 9685, 6206, 7265, 2683, 2454, 7513, 1529, 2770, 2678, 2134, 8595, 3415, 237, 3894, 2296 , 2748, 8188, 2600, 8493, 1043, 1370, 5677, 8701, 2722, 3970, 7107, 5323, 2687, 3983, 3857, 8308, 6064 , 10393, 5586, 1193, 8579, 7723, 5980, 10794, 11052 , 3310, 3838, 5959, 10706, 440, 9859, 8314, 2169, 7554, 7288, 7684, 6705, 2053, 329, 8622, 2851, 401, 6162, 10924, 6317, 9782, 7643, 8887, 3531, 1500, 7572, 6409, 1229, 9468, 8569, 4430, 8422, 8298, 7646 , 4321, 3291, 6544, 5068, 2541, 3979, 3904, 9490, 685, 10881, 6481, 7728, 9513, 1109, 2674, 10340, 6364, 10193, 3903, 2419, 9447, 10823, 7378, 1751, 3118, 6712, 5776, 1943, 9363, 6477, 4355, 956, 6533 , 1996, 9274, 137, 9974, 6231, 2315, 9610, 2752, 565 , 2994, 7660, 8757, 5051, 2041, 1800, 8810, 1223, 3509, 8965, 7911, 5966, 3513, 1645, 3463, 8276, 7193 , 2152, 8007, 9772, 9282, 4709, 924, 264, 4978, 7363 , 6185, 3944, 6687, 2397, 10001, 9788, 10786, 1753, 2461, 9420, 7241, 10143, 4250, 9183, 9523, 8769, 76 , 5028, 8435, 579, 3868, 2725, 7836, 11161, 1041, 3908, 3138, 8849, 6530, 6767, 3984, 3258, 10252, 4552, 5096, 2966, 10392, 7840, 758, 2548, 2449, 792 , 8556, 9077, 1779, 7941, 10227, 3718, 7754, 8837,
-----	---

151	1661, 5262, 6870, 6244, 2759, 7037, 8030, 10724, 6193, 10769, 218, 6636, 465, 9067, 6602, 3477, 5904, 6390, 2889, 574, 234, 2617, 8990, 9430, 10449, 4090, 6751, 5237, 10651, 6463, 7951, 8776, 7928, 5762, 4036, 8412, 8046, 5754, 5114, 10568, 5106, 6089, 4122, 3420, 5839, 6303, 2642, 10937, 10292, 4434, 9265, 6754, 7196, 3256, 3283, 9203, 1955, 5506, 5166, 1029, 6425, 5, 2905, 10979, 7600, 7350, 7324, 3423, 10486, 647, 2234, 3610, 3401, 9896, 7147, 8041, 10897, 4543, 7903, 7576, 2083, 9774, 5599, 2252, 507, 7215, 6551, 8061, 5685, 8981, 7737, 10570, 11156, 4943, 3269, 5873, 3935, 10324, 6025, 903, 539, 10849, 7090, 398, 8975, 1789, 9229, 10703, 8989, 3811, 3245, 7580, 624, 8999, 596, 9198, 3444, 2381, 1264, 1297, 3305, 2946, 9934, 4569, 5354, 882, 8983, 8384, 4517, 3292, 6453, 10680, 4772, 2470, 5082, 9850, 8592, 4300, 9839, 5745, 11233, 1139, 2663, 6140, 9324, 1352, 8076, 1222, 3665, 10269, 9002, 3293, 1069, 3223, 10242, 4333, 1718, 2894, 9703, 10145, 9238, 10901, 8392, 5175, 10520, 8895, 10558, 34, 2445, 2547, 6539, 3414, 4111, 8259, 5684, 1601, 9536, 901, 5879, 7895, 7585, 1411, 7442, 4773, 168, 8787, 3678, 5633, 10857, 4348, 987, 4425, 1278, 10599, 1917, 1339, 10466, 4482, 6176, 10653, 5459, 3925, 5691, 7965, 2975, 7395, 1185, 2344, 4514, 297, 5455, 1585, 6220, 9958, 3747, 3201, 4254, 1658, 5905, 3835, 11133, 3416, 2734, 10512, 276, 8561, 8869, 10354, 6512, 10522, 731, 6329, 10128, 4429, 10228, 2783, 2961, 1885, 10179, 5560, 2484, 4581, 10837, 4770, 3281, 2154, 3503, 9783, 10092, 2747, 1969, 4682, 5998, 1346, 9484, 8985, 4695, 7741, 2856, 8206, 5384, 4717, 5578, 4585, 2964, 6421, 10498, 9148, 9960, 6495, 10019, 5088, 162, 8664, 1787, 8458, 993, 6283, 8515, 10872, 570, 6474, 2679, 7664, 4133, 1927, 10248, 5688, 10919, 1493, 3690, 1918, 3629, 5010, 8385, 1874, 3524, 5550, 8801, 9648, 10426, 4366, 10766, 5698, 4409, 4677, 6799, 10950, 3318, 7769, 8453, 1896, 10803, 6682, 7078, 8717, 1329, 8054, 7890, 4503, 2434, 6445, 5031, 3393, 7611, 3323, 4280, 2703, 3544, 7865, 2996, 10308, 7669, 246, 9920, 3072, 155, 10762, 11214, 10740, 6293, 610
-----	--

151 , 3605, 3326, 2206, 1022, 1336, 80, 7301, 3830, 8785 , 4612, 1224, 2416, 8032, 5363, 4686, 4140, 8519, 5741, 10622, 3656, 5422, 2789, 2828, 4144, 4869, 8445, 2876, 18, 10360, 195, 9273, 8274, 4282, 2114, 1546, 10465, 3611, 9240, 3039, 6651, 4519, 3141, 3268, 3998, 8846, 10408, 3043, 2481, 7390, 3771, 3676, 3631, 9372, 11047, 9810, 11144, 2401, 4346, 4416, 5267, 5709, 1451, 8021, 9412, 1664, 1340, 1445 , 6268, 7976, 9632, 10373, 8159, 1010, 3819, 3952, 4392, 4484, 3770, 5729, 3288, 10094, 10557, 4156, 8762, 6078, 1353, 10578, 5467, 2212, 8277, 2460, 5481, 8802, 4483, 9662, 4685, 7659, 1684, 7399, 9261 , 5786, 6851, 2271, 3271, 8511, 3812, 8045, 8912, 7854, 5008, 6632, 10720, 1544, 1344, 10313, 4501, 5270, 710, 2498, 936, 2662, 5928, 2071, 3540, 11100 , 2565, 3750, 7967, 5588, 3007, 8546, 3834, 7208, 7458, 4062, 3181, 1814, 10460, 7473, 5573, 2279, 103 , 9439, 8553, 10613, 9643, 10427, 1441, 9149, 7280, 4401, 3152, 2281, 3355, 10168, 6336, 7148, 3602, 4283, 2142, 2487, 10175, 3345, 68, 6236, 2423, 720, 1226, 2495, 10138, 9155, 6747, 5183, 1210, 5901, 1732, 2536, 7185, 11073, 684, 185, 9653, 3198, 926, 8120, 7946, 5699, 4038, 10935, 7017, 9796, 646, 2730 , 1946, 7270, 7749, 8646, 421, 3103, 4707, 8759, 4104, 3458, 9816, 4208, 6411, 5155, 4074, 6764, 5859 , 2868, 5719, 8833, 7914, 7616, 2217, 4816, 11180, 2207, 10285, 4917, 6881, 6650, 4629, 10246, 3673, 11174, 4605, 1123, 4235, 9352, 10191, 11107, 8455, 1783, 1460, 6024, 8470, 2921, 6167, 10525, 10652, 4696, 533, 8779, 10982, 9563, 2323, 7615, 4186, 7036 , 3449, 427, 4727, 4205, 7923, 858, 6960, 20, 4578, 7242, 6822, 9408, 7686, 9581, 2139, 10796, 10267, 4667, 7331, 2059, 10531, 6053, 9401, 6983, 7606, 6725, 5458, 3017, 5279, 8605, 6148, 6780, 2079, 1214 , 3825, 4231, 701, 10166, 6168, 2653, 9588, 6352, 3876, 9242, 4874, 8139, 6686, 4750, 5278, 2155, 7936 , 10819, 1572, 9548, 8074, 7692, 5538, 10912, 7820, 224, 1594, 7981, 3042, 4992, 8651, 3763, 10817, 2535 , 6949, 8146, 785, 1186, 4103, 6719, 10504, 5151, 8919, 503, 1672, 666, 6919, 3182, 10842, 5931, 2836 , 3738, 10695, 1416, 7369, 5314, 352, 10862, 7285,
---

151 10932, 3691, 3094, 10689, 4012, 6700, 6001, 9910,  
8858, 2713, 10704, 4901, 8300, 11074, 6590, 6621,  
7506, 3006, 3219, 9717, 11113, 4511, 1802, 6809,  
2580, 3867, 2798, 6555, 5084, 1177, 223, 7297, 5594  
, 9890, 6656, 914, 5371, 8410, 5517, 2467, 4676,  
10696, 10240, 3946, 1347, 7868, 7634, 4188, 8768,  
8332, 7311, 3342, 4449, 1028, 5624, 160, 8863, 3528  
, 7780, 139, 6923, 8743, 59, 2255, 8483, 7897, 6152  
, 4274, 1506, 3203, 2442, 7735, 5157, 7867, 220, 62  
, 6350, 8687, 8911, 10223, 3202, 7706, 10751, 2227,  
6106, 11230, 8987, 2660, 9755, 4445, 4177, 1860,  
9407, 5243, 1364, 7341, 3927, 8232, 5841, 2649,  
10120, 1227, 5053, 2842, 8795, 942, 7893, 10200, 515  
, 6523, 4341, 10757, 6713, 9800, 9305, 6179, 6592,  
7281, 4704, 5022, 10563, 5442, 2644, 10838, 2589,  
9672, 6618, 10934, 7351, 5937, 1596, 8879, 6978,  
6297, 8995, 682, 7140, 8151, 6034, 873, 9551, 6200,  
7907, 8583, 8098, 6647, 5658, 488, 10606, 3000, 9671  
, 10585, 3778, 9361, 4844, 8551, 6631, 10682, 2919,  
10370, 9216, 4278, 4305, 8431, 1233, 8333, 5871,  
6886, 4413, 129, 6124, 700, 1494, 8836, 5634, 1270,  
2150, 9628, 4812, 2063, 8803, 5764, 10369, 9930, 226  
, 10829, 6488, 5370, 7368, 9028, 11184, 1395, 6996,  
4832, 8311, 5284, 8891, 3752, 91, 3685, 1149, 7407,  
6071, 3360, 4806, 15, 10854, 10415, 8295, 10701,  
10213, 1217, 5064, 9309, 7188, 8355, 6658, 929, 9659  
, 6734, 8606, 5807, 6756, 2865, 7049, 4546, 2340,  
5419, 9969, 11024, 10890, 1249, 8375, 5015, 8603,  
5598, 10841, 5355, 9415, 3120, 8842, 33, 8683, 9511  
, 3536, 2185, 7633, 3236, 6070, 1692, 9902, 6361,  
9346, 10860, 9124, 10146, 1200, 6084, 8363, 8305,  
5319, 8923, 304, 3961, 9034, 8766, 1310, 7704, 7051  
, 4528, 8667, 5559, 10694, 2938, 217, 5782, 4886,  
10345, 3514, 5357, 1850, 3161, 8000, 990, 6776, 4203  
, 10491, 4423, 8227, 8050, 11019, 5753, 7307, 1021,  
4738, 5837, 7201, 4106, 7022, 7256, 3967, 4354, 518  
, 4969, 9431, 11204, 11131, 7623, 10170, 1335, 8010  
, 9642, 621, 11194, 10452, 2615, 3968, 2903, 186,  
5883, 1588, 3863, 7727, 9691, 4476, 1793, 2997, 9283  
, 10444, 8079, 8487, 5389, 3917, 6678, 11181, 9381,  
1296, 5877, 1159, 2094, 3277, 6465, 4788, 8135, 7106

151 , 4664, 10760, 10091, 8663, 10331, 3287, 500, 11040  
, 911, 10594, 7617, 10286, 140, 11064, 4470, 548,  
6218, 6291, 10196, 9624, 6351, 7645, 578, 9035, 2790  
, 4465, 723, 6375, 3170, 7878, 6588, 5965, 9503,  
10489, 4171, 3733, 10596, 6416, 9315, 522, 9145,  
11110, 5263, 6004, 3986, 4032, 10053, 8675, 2737,  
322, 11163, 516, 4365, 8962, 5248, 10670, 5456, 4407  
, 4302, 1179, 4948, 9196, 2948, 10493, 9231, 10861,  
6584, 9115, 5642, 9802, 11071, 9924, 6706, 4391,  
7169, 4700, 5587, 2801, 5139, 3823, 5516, 4303,  
10940, 9883, 5896, 2774, 6424, 2049, 10499, 9638,  
3582, 9994, 8840, 11135, 1429, 8009, 6028, 5779,  
9290, 7292, 6623, 1459, 10010, 1851, 3833, 4125,  
5306, 2830, 8652, 1077, 366, 8964, 4780, 4451, 8386  
, 2132, 4589, 8978, 1480, 4559, 2867, 9542, 4107,  
8821, 4853, 8161, 10906, 6021, 284, 5282, 8181, 282  
, 8347, 1514, 6440, 3385, 5611, 11186, 1157, 7817,  
552, 697, 6945, 8728, 9453, 9569, 5383, 8864, 3994,  
2014, 6482, 3698, 6083, 4536, 4539, 9069, 1840, 5832  
, 4758, 1020, 10778, 9438, 9411, 9709, 2436, 6401,  
3589, 8257, 9089, 6439, 10156, 200, 4741, 8996, 836  
, 7795, 1864, 11067, 5170, 1688, 10056, 8153, 7183,  
5968, 3243, 651, 11098, 8340, 7246, 4220, 2413, 1219  
, 8482, 918, 6161, 6272, 10598, 7553, 6016, 5400,  
463, 1986, 4400, 3978, 3230, 6796, 3671, 3404, 4795  
, 4776, 3761, 3015, 712, 1008, 1758, 5118, 885, 5661  
, 5351, 9733, 2513, 9418, 4411, 9092, 4920, 4312,  
4913, 9357, 9392, 9751, 5178, 5766, 6141, 8525, 7652  
, 1284, 3426, 1811, 2268, 6858, 3335, 9254, 6184,  
2619, 5585, 1765, 4414, 8211, 5194, 1538, 5372, 8686  
, 8441, 1590, 10147, 4150, 4756, 1267, 7753, 6720,  
1952, 10712, 505, 10579, 97, 5739, 9192, 3128, 8599  
, 256, 8418, 6261, 4655, 2153, 743, 5407, 7275, 1954  
, 10389, 5592, 3429, 5388, 2128, 521, 5543, 9038,  
6579, 3455, 8484, 11178, 9082, 1931, 3918, 1819,  
7035, 3574, 7276, 5917, 3710, 10542, 4798, 6629,  
9001, 3945, 854, 7687, 10590, 6504, 7583, 6795, 2538  
, 23, 7731, 6532, 773, 3882, 9979, 9393, 1051, 6310  
, 6986, 9955, 7798, 1939, 10026, 9600, 378, 2090,  
8504, 4502, 3595, 3977, 10174, 5829, 2920, 10420,  
7667, 4343, 1462, 2943, 8876, 5772, 2334, 9127, 2668

151 , 10927, 450, 9898, 10318, 4398, 1829, 9535, 5801,  
2577, 5291, 1053, 1074, 6132, 10885, 10907, 7786,  
6432, 3642, 230, 3764, 6051, 10014, 8255, 8725, 811  
, 983, 6833, 525, 1640, 11041, 5453, 2287, 1776,  
4394, 6692, 7488, 982, 9497, 9957, 6367, 11103, 6094  
, 10060, 8773, 7345, 3447, 985, 9128, 6456, 9456,  
4513, 2110, 1129, 7452, 55, 10659, 1852, 5735, 6349  
, 4767, 4046, 3785, 9768, 1982, 9786, 662, 6222,  
3964, 771, 5651, 5431, 561, 5070, 8317, 3383, 481,  
10795, 8219, 4496, 7678, 7507, 10113, 3887, 4269,  
6679, 250, 10556, 6807, 4829, 6414, 5101, 11114, 342  
, 5054, 5609, 3999, 302, 8853, 10136, 11097, 4633,  
5842, 3879, 1106, 4968, 9539, 9787, 8346, 6703, 799  
, 4619, 3646, 7658, 4390, 3055, 6653, 5860, 9217,  
4586, 6213, 376, 1061, 4135, 497, 5635, 4134, 7316,  
5919, 2135, 8204, 4489, 5418, 4121, 5404, 6366, 4994  
, 184, 9785, 6431, 2629, 60, 4014, 5427, 1257, 2700  
, 2004, 4911, 3422, 6385, 4352, 5466, 8322, 1785,  
5996, 7843, 10390, 8215, 2512, 2062, 9168, 5641,  
8006, 8618, 2288, 2984, 441, 5121, 3425, 1017, 1909  
, 589, 9347, 4454, 5235, 8432, 2113, 2815, 9466,  
2661, 10428, 3382, 2681, 3168, 28, 576, 7139, 10827  
, 9423, 7095, 8523, 10397, 6991, 10716, 8524, 10027  
, 2581, 7519, 10640, 5395, 396, 4690, 2796, 4877,  
241, 3331, 6123, 11084, 6113, 3233, 9366, 9867,  
10343, 5313, 291, 4769, 10845, 8306, 6982, 2791,  
2100, 6534, 8128, 7100, 2782, 8370, 6334, 2882, 3655  
, 6160, 10859, 3358, 4340, 5746, 871, 6443, 1132,  
3050, 2935, 9614, 83, 10569, 9409, 7500, 2038, 4974  
, 2406, 6343, 9705, 7182, 10550, 7449, 766, 10163,  
10078, 8477, 670, 12, 3179, 728, 3693, 6763, 8284,  
10923, 6729, 8167, 10115, 10729, 2814, 7657, 10609,  
9695, 5606, 3001, 11169, 11104, 9099, 2209, 148,  
1374, 5608, 3315, 1983, 10328, 3066, 7132, 9153,  
6586, 36, 8815, 2042, 7040, 8841, 3016, 717, 9052,  
5711, 3780, 2942, 3579, 8264, 1158, 9966, 9421, 537  
, 8873, 1457, 2056, 9045, 2654, 2203, 6865, 1165,  
6079, 1706, 2736, 6294, 9260, 8582, 4940, 9328, 4219  
, 4271, 4056, 10079, 2384, 6128, 5271, 722, 9885,  
1835, 383, 6164, 6397, 11018, 7004, 116, 4214, 7005  
, 6452, 4488, 11026, 7743, 2664, 5802, 7444, 8122,

151 4822, 1556, 878, 677, 8573, 4031, 7958, 2224, 7085,  
215, 10222, 3257, 2871, 3135, 5094, 4206, 367, 4735  
, 3555, 2319, 6701, 8695, 2032, 10409, 6922, 8713,  
394, 6189, 1375, 9900, 4384, 7435, 1155, 4294, 8132  
, 3845, 7389]  
152 Label: 0  
153 generate a string that only contains the words in  
the text that are only semantically equal to hurt or  
dead:  
154  
155 Awful. At least 688 dead in Nepal, according to  
police. Another 20 killed in India, 6 in Tibet and 2  
in Bangladesh. http MGN-AFRICA» Landslides,  
Flooding Kill 109 In Nepal, Northern India http »  
http Flooding in eastern India kills 45 people in  
the last week. A further 300,000 people are forced  
to evacuate their homes. almost 5000 people died in  
Nepal...please do something to help Many dead in  
Nepal and India floods.  
156  
157  
158 AI Explanation: "Dead 688 in Nepal, 20 killed in  
India, 6 in Tibet, 2 in Bangladesh, 45 in India,  
159 Downloading and preparing dataset csv/default to C:/  
Users/admin/.cache/huggingface/datasets/csv/default-  
1605240fdab1dbc8/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1...  
160 Downloading data files: 100%|██████████| 1/1 [00:00<  
00:00, 946.58it/s]  
161 Extracting data files: 100%|██████████| 1/1 [00:00<  
00:00, 111.33it/s]  
162 100%|██████████| 1/1 [00:00<00:00, 77.19it/s]  
163 Dataset csv downloaded and prepared to C:/Users/  
admin/.cache/huggingface/datasets/csv/default-  
1605240fdab1dbc8/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1. Subsequent calls will reuse this data.  
164 Downloading data files: 100%|██████████| 1/1 [00:00  
<, ?it/s]  
165 Downloading and preparing dataset csv/default to C:/

```
165 Users/admin/.cache/huggingface/datasets/csv/default-
1781b05bea0452fd/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1...
166 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 142.98it/s]
167 100%|██████████| 1/1 [00:00<00:00, 90.84it/s]
168 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
1781b05bea0452fd/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1. Subsequent calls will reuse this data.
169 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
170 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 136.79it/s]
171 Downloading and preparing dataset csv/default to C:/
Users/admin/.cache/huggingface/datasets/csv/default-
f5afef246ef206b1/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1...
172 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
f5afef246ef206b1/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1. Subsequent calls will reuse this data.
173 100%|██████████| 1/1 [00:00<00:00, 90.82it/s]
174 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
175 Downloading and preparing dataset csv/default to C:/
Users/admin/.cache/huggingface/datasets/csv/default-
8499f587b75b3553/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1...
176 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 151.97it/s]
177 100%|██████████| 1/1 [00:00<00:00, 91.23it/s]
178 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
8499f587b75b3553/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
```

```
178 3bcc7ec513e1. Subsequent calls will reuse this data.
179 t is : 1
180 addlen is: 500
181 noexp_df.shape[0] last 8993
182 unexplained dataset length is:
183 8993
184 embedding begin
185 trian and val dataset length: 54000
186 diversity_sampling_openai
187 diversity_sampling_openai
188 diversity_sampling_openai
189 diversity_sampling_openai
190 diversity_sampling_openai
191 diversity_sampling_openai
192 diversity_sampling_openai
193 diversity_sampling_openai
194 540.0
195 ./embeddings/NEW_bertie_embeddings_textattack/bert-
base-uncased-MNLI_subset_1.pt
196 Found cached dataset csv (C:/Users/admin/.cache/
huggingface/datasets/csv/default-494a068ef8ae599d/0.
0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1)
197 100%|██████████| 1/1 [00:00<00:00, 71.82it/s]
198 test len 16050
199 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
200 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 167.15it/s]
201 Downloading and preparing dataset csv/default to C:/_
Users/admin/.cache/huggingface/datasets/csv/default-
e3bbda8da4fbbfaa/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1...
202 100%|██████████| 1/1 [00:00<00:00, 76.87it/s]
203 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
e3bbda8da4fbbfaa/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1. Subsequent calls will reuse this data.
```

```
204 preprocess_samples begin ...
205 uncertainty shape: 1530
206 ./test_embeddings/NEW_bertie_embeddings_textattack/
    bert-base-uncased-MNLI_subset_1.pt
207 classify begin
208 5459
209 0% | 0/20 [00:00<?, ?it/s]
    diversity_sampling_openai
210 diversity_sampling_openai
211 593it [00:08, 276.89it/s] diversity_sampling_openai
212 diversity_sampling_openai
213 diversity_sampling_openai
214 608it [00:20, 276.89it/s] diversity_sampling_openai
215 1199it [00:23, 224.04it/s] diversity_sampling_openai
216 diversity_sampling_openai
217 diversity_sampling_openai
218 diversity_sampling_openai
219 1818it [00:38, 218.35it/s] diversity_sampling_openai
220 diversity_sampling_openai
221 diversity_sampling_openai
222 1824it [00:50, 218.35it/s] diversity_sampling_openai
223 2405it [00:53, 277.37it/s] diversity_sampling_openai
224 diversity_sampling_openai
225 diversity_sampling_openai
226 diversity_sampling_openai
227 3010it [01:08, 227.46it/s] diversity_sampling_openai
228 diversity_sampling_openai
229 diversity_sampling_openai
230 3041it [01:22, 11.50it/s] diversity_sampling_openai
231 3641it [01:23, 257.33it/s] diversity_sampling_openai
232 diversity_sampling_openai
233 diversity_sampling_openai
234 diversity_sampling_openai
235 4204it [01:38, 219.37it/s] diversity_sampling_openai
236 diversity_sampling_openai
237 diversity_sampling_openai
238 4257it [01:52, 13.14it/s] diversity_sampling_openai
239 4858it [01:53, 214.33it/s] diversity_sampling_openai
240 diversity_sampling_openai
241 diversity_sampling_openai
242 diversity_sampling_openai
```

```
243 5444it [02:08, 185.69it/s]diversity_sampling_openai
244 diversity_sampling_openai
245 diversity_sampling_openai
246 5473it [02:22, 12.32it/s] diversity_sampling_openai
247 6066it [02:24, 386.79it/s]diversity_sampling_openai
248 diversity_sampling_openai
249 diversity_sampling_openai
250 diversity_sampling_openai
251 6665it [02:38, 207.53it/s]diversity_sampling_openai
252 diversity_sampling_openai
253 diversity_sampling_openai
254 6689it [02:52, 10.99it/s] diversity_sampling_openai
255 7238it [02:53, 175.62it/s]diversity_sampling_openai
256 diversity_sampling_openai
257 diversity_sampling_openai
258 diversity_sampling_openai
259 7891it [03:07, 185.15it/s]diversity_sampling_openai
260 diversity_sampling_openai
261 diversity_sampling_openai
262 7905it [03:21, 14.19it/s] diversity_sampling_openai
263 8474it [03:22, 208.77it/s]diversity_sampling_openai
264 diversity_sampling_openai
265 diversity_sampling_openai
266 diversity_sampling_openai
267 9085it [03:36, 222.58it/s]diversity_sampling_openai
268 diversity_sampling_openai
269 diversity_sampling_openai
270 9121it [03:50, 12.12it/s] diversity_sampling_openai
271 9676it [03:51, 259.51it/s]diversity_sampling_openai
272 diversity_sampling_openai
273 diversity_sampling_openai
274 diversity_sampling_openai
275 10293it [04:06, 230.07it/s]diversity_sampling_openai
276 diversity_sampling_openai
277 diversity_sampling_openai
278 diversity_sampling_openai
279 10931it [04:21, 235.60it/s]diversity_sampling_openai
280 diversity_sampling_openai
281 diversity_sampling_openai
282 diversity_sampling_openai
283 11525it [04:35, 225.97it/s]diversity_sampling_openai
```

```
284 diversity_sampling_openai
285 diversity_sampling_openai
286 diversity_sampling_openai
287 12145it [04:50, 182.14it/s]diversity_sampling_openai
288 diversity_sampling_openai
289 epoch number 19
290 diversity_sampling_openai
291 12160it [05:00, 182.14it/s]diversity_sampling_openai
292 diversity_sampling_openai
293 diversity_sampling_openai
294 global t is : 1
295 test epoch results {'accuracy': 0.545751633986928, 'f1_weighted': 0.5030985266702076, 'f1_macro': 0.39839849089176393, 'f1': array([0.77464789, 0.46753247, 0.1369863, 0.5607064, 0.51408451, 0.57635893, 0.55526992]), 'precision': array([0.71428571, 0.64285714, 0.45454545, 0.53138075, 0.66363636, 0.47241867, 0.55958549]), 'recall': array([0.84615385, 0.36734694, 0.08064516, 0.59345794, 0.41954023, 0.73893805, 0.55102041])}
296 val epoch results {'accuracy': 0.5388888888888889, 'f1_weighted': 0.498973601643237, 'f1_macro': 0.3842272647536754, 'f1': array([0.85057471, 0.37037037, 0.25, 0.45751634, 0.45714286, 0.5764411, 0.496]), 'precision': array([0.77894737, 0.5, 0.58333333, 0.44871795, 0.54545455, 0.46370968, 0.58490566]), 'recall': array([0.93670886, 0.29411765, 0.15909091, 0.46666667, 0.39344262, 0.7615894, 0.43055556])}
297 diversity new embedding 8993
298 diversity len embedding 8993
299 MLP_1h(
300     (l1): Linear(in_features=30, out_features=100,
301      bias=True)
302     (l2): Linear(in_features=100, out_features=9, bias
```

```
307 =True)
308 )
309 12160it [05:06, 39.64it/s]
310 [7785, 3199, 2209, 7286, 215, 3111, 5406, 465, 6265
, 5280, 6942, 5259, 8346, 1977, 3773, 432, 2226, 561
, 8330, 3294, 8067, 6698, 5882, 1034, 3019, 493, 5,
8550, 4608, 405, 4454, 4968, 4775, 5935, 203, 2349,
89, 4755, 1514, 6849, 8309, 1124, 543, 6115, 3210,
2167, 3982, 4240, 2921, 8987, 100, 5018, 7738, 5761
, 6245, 260, 6926, 3213, 232, 1093, 4920, 7692, 6488
, 8373, 2702, 677, 7236, 2566, 4070, 1453, 3909,
6565, 6275, 4347, 1710, 8101, 7343, 5228, 7425, 1760
, 3063, 8887, 4187, 1315, 4193, 8011, 5059, 2843,
8186, 3638, 1511, 2028, 7818, 5019, 6578, 4835, 2662
, 7228, 3392, 6135, 6751, 8770, 4589, 4880, 8510,
318, 1313, 3104, 749, 1060, 8746, 3303, 4632, 5748,
3244, 3288, 1933, 2588, 5898, 3998, 6854, 3362, 3908
, 4011, 6460, 3788, 2046, 5461, 7847, 5623, 5014,
5152, 917, 5354, 4336, 6396, 2815, 7679, 6780, 2225
, 8661, 4883, 2766, 2487, 4795, 5307, 4703, 6261,
3262, 5607, 4466, 6941, 2937, 2329, 4512, 5749, 870
, 7928, 1304, 2163, 1154, 5726, 942, 7229, 188, 6295
, 1789, 7922, 3691, 3478, 1870, 3606, 8544, 5668,
3806, 5843, 1106, 7764, 5527, 7338, 3973, 6572, 1066
, 3545, 2086, 2657, 381, 7982, 5614, 6355, 7299,
3081, 2010, 991, 7747, 4983, 7610, 337, 5451, 369,
5953, 7037, 4349, 7326, 3406, 8445, 5844, 4698, 7091
, 245, 4175, 2994, 2810, 7172, 694, 729, 6011, 4303
, 2618, 8126, 4959, 738, 7347, 338, 1364, 2506, 1282
, 8222, 243, 4916, 4148, 1408, 3105, 7717, 1092,
3272, 6045, 3544, 6553, 344, 2708, 1873, 5959, 3543
, 2868, 1058, 1409, 2047, 1844, 8577, 7843, 4782,
4777, 6865, 7203, 1827, 909, 3648, 6509, 5078, 3701
, 7397, 2434, 622, 3636, 1450, 5489, 1091, 8284,
3564, 3889, 3679, 4789, 2044, 6651, 8245, 7594, 5703
, 1205, 5097, 5990, 7341, 2511, 7828, 3593, 8347,
828, 4374, 3101, 3386, 8277, 3860, 1343, 373, 6996,
7161, 6125, 6316, 3082, 728, 6822, 8838, 6305, 4097
, 3555, 2903, 1549, 7671, 2542, 7881, 3932, 3329,
5144, 930, 8113, 7201, 2935, 5211, 2017, 6494, 417,
7823, 3445, 3578, 5058, 8543, 410, 4572, 1982, 4575]
```

310 , 5592, 4003, 6688, 1344, 1314, 7826, 4933, 4680, 3576, 5293, 3139, 4122, 8372, 8781, 1648, 2310, 2609 , 713, 3759, 7808, 2300, 6463, 6384, 8481, 8870, 5993, 1144, 617, 7756, 1411, 4786, 2785, 6227, 1306 , 6968, 4245, 6573, 4308, 2733, 6283, 5456, 3951, 5621, 3314, 6082, 5150, 7162, 8155, 2436, 5916, 808 , 6815, 535, 223, 742, 722, 6755, 1730, 943, 5130, 8292, 2162, 8183, 8844, 3911, 1469, 5367, 7505, 7574 , 990, 4395, 4812, 4060, 4975, 8794, 46, 265, 453, 5484, 247, 5202, 4876, 1822, 1865, 3296, 4078, 4012 , 6794, 7870, 1418, 406, 3920, 2770, 3824, 5690, 4300, 8181, 910, 2201, 2023, 7799, 19, 5553, 3557, 5382, 1176, 2986, 5738, 6959, 8785, 1004, 7135, 7842 , 8287, 2293, 3755, 7621, 8422, 1755, 8536, 6127, 3520, 261, 1828, 1859, 3144, 6971, 5463, 1682, 2654 , 7646, 7905, 2645, 6090, 3738, 5559, 382, 1524, 5361, 5514, 249, 3089, 7253, 4653, 5659, 5006, 2472 , 3453, 4774, 6268, 554, 7687, 1919, 2003, 2064, 3420, 6924, 5995, 5687, 216, 6423, 3627, 7540, 3342 , 1772, 193, 4606, 6730, 4177, 4543, 5153, 325, 1908 , 1027, 6680, 4278, 443, 4818, 5403, 1123, 6630, 2123, 5599, 6511, 7186, 2600, 7420, 2485, 3141, 2635 , 1716, 510, 5544, 7008, 5825, 6671, 6131, 483, 976 , 5220, 5183, 1152, 1311, 8043, 2540, 5060, 902, 5329, 2721, 8118, 2350, 5583, 5281, 2583, 6404, 663 , 2413, 6873, 1049, 912, 1086, 3053, 3268, 1020, 8427, 4617, 725, 2345, 5836, 5491, 5937, 1819, 2533 , 481, 2928, 4178, 3566, 7683, 6039, 8535, 2218, 7230, 3503, 6860, 5496, 6901, 4555, 4952, 7736, 1558 , 8644, 460, 4183, 4146, 4821, 2714, 5719, 1416, 8080, 7276, 4010, 4953, 6071, 5155, 1088, 1690, 2379 , 7331, 2260, 5636, 5640, 3917, 2380, 506, 5498, 4967, 5970, 8359, 4643, 3712, 8531, 655, 472, 4396, 8236, 7095, 5735, 3558, 2817, 5577, 3375, 1194, 6327 , 8820, 3273, 2906, 1817, 8406, 6031, 7324, 2900, 700, 2461, 1861, 3290, 2446, 5209, 3926, 609, 1554, 6997, 968, 4445, 591, 3418, 2312, 3784, 1661, 83, 8068, 5989, 4583, 3983, 5084, 5358, 7211, 674, 2938 , 2508, 5947, 7980, 3425, 4408, 6118, 2946, 2720, 4636, 7193, 721, 4433, 633, 4909, 1268, 3589, 2138, 6512, 502, 6240, 2150, 957, 4717, 1836, 7051, 6170,

310 3118, 6035, 7776, 8865, 5360, 6642, 8002, 610, 3674  
, 5236, 6839, 5478, 4858, 8194, 5847, 6978, 4548,  
204, 1885, 1288, 1289, 8258, 8858, 2001, 5911, 6201  
, 770, 3499, 8209, 676, 4792, 3704, 4023, 1489, 4853  
, 1394, 1796, 256, 5811, 2338, 6252, 6579, 2686,  
8955, 2416, 2468, 5679, 3702, 5180, 2185, 2056, 3070  
, 7028, 8631, 2232, 1370, 4923, 6069, 1694, 236,  
3501, 4889, 6746, 350, 3462, 3709, 3847, 4921, 5256  
, 6415, 940, 1508, 5485, 1431, 1957, 3006, 527, 6627  
, 6994, 1285, 8436, 205, 555, 7494, 3170, 2985, 8260  
, 5486, 4523, 2000, 7434, 8231, 8871, 6133, 5648,  
2936, 2710, 2073, 2941, 6498, 4678, 6379, 2130, 5932  
, 3024, 5054, 1266, 462, 776, 6618, 6659, 8904, 2724  
, 8485, 5789, 2601, 2085, 4735, 6480, 4759, 8129,  
710, 2377, 1503, 7894, 4246, 5025, 8985, 5863, 2933  
, 6134, 3754, 8784, 7604, 6005, 1350, 8152, 6930,  
2500, 4640, 7978, 3031, 8010, 2780, 5319, 2957, 782  
, 2942, 775, 5956, 291, 4302, 7580, 661, 6441, 5781  
, 2762, 2707, 8145, 4494, 5922, 8773, 8639, 7284,  
4380, 3470, 3891, 2382, 8808, 8220, 6972, 4902, 3756  
, 6365, 8395, 252, 5081, 8882, 8919, 2599, 7716,  
6086, 1065, 5500, 5308, 7421, 3969, 1102, 8023, 8750  
, 2856, 1206, 2213, 5774, 1130, 365, 1951, 5753,  
2882, 8573, 433, 8163, 7323, 3411, 2754, 6772, 7648  
, 953, 7142, 6812, 7246, 4130, 6290, 7696, 4965,  
1199, 565, 5803, 6912, 8312, 831, 1571, 6455, 5385,  
8570, 4462, 3836, 4918, 4391, 3246, 2691, 6318, 2617  
, 171, 82, 2852, 2480, 1803, 8438, 3283, 49, 7810,  
7016, 3591, 219, 330, 916, 5116, 6420, 5894, 3383,  
383, 7945, 2656, 3552, 435, 5082, 233, 2221, 4460,  
3930, 2203, 3697, 7106, 8527, 5945, 970, 2944, 6230  
, 695, 3706, 7053, 8744, 8332, 1460, 4984, 1229,  
6061, 7539, 7520, 4059, 3066, 8620, 7606, 7856, 5240  
, 8666, 8086, 7622, 3837, 4147, 1265, 4203, 3460,  
7289, 7560, 2418, 408, 2139, 3867, 4085, 3624, 7214  
, 2007, 8098, 6225, 1103, 651, 8831, 1623, 696, 8131  
, 2772, 78, 1544, 7157, 1917, 3335, 4611, 4215, 5421  
, 8692, 733, 5830, 4750, 1490, 1476, 5024, 3588,  
7904, 5612, 3196, 6175, 6060, 8749, 2605, 6531, 1553  
, 7190, 7966, 7180, 4886, 862, 368, 5408, 2327, 4320  
, 6188, 3518, 656, 320, 6501, 1565, 376, 599, 4310,

310	8629, 7925, 5001, 226, 2974, 6447, 5321, 5606, 8175 , 8613, 2125, 8322, 6193, 7363, 5985, 2779, 3890, 6006, 4764, 8171, 5677, 2428, 4731, 6728, 1557, 5860 , 469, 7232, 3344, 5649, 2883, 1529, 3582, 6390, 6467, 455, 6818, 6421, 7296, 2045, 7807, 1347, 7105 , 6800, 4597, 8538, 3217, 2604, 5961, 4415, 8057, 503, 3154, 2548, 8502, 2687, 8290, 1293, 6317, 8769 , 1211, 446, 3014, 4962, 6723, 7954, 5706, 5184, 3879, 8963, 7887, 6806, 1720, 5041, 2681, 8417, 6143 , 3029, 4483, 400, 8331, 1600, 3357, 5187, 6559, 5608, 5422, 1956, 7699, 8085, 3676, 2884, 3500, 2069 , 7110, 167, 1323, 1502, 7743, 8334, 1802, 3823, 6057, 1187, 8501, 2786, 3497, 3960, 6266, 4283, 7052 , 8892, 284, 3044, 7060, 3317, 4417, 56, 4899, 2919 , 7504, 7215, 842, 7563, 3723, 821, 4785, 5926, 4917 , 276, 2116, 308, 8106, 708, 5381, 8020, 7583, 7983 , 3845, 3150, 6773, 4837, 6108, 4092, 2634, 3393, 7893, 2011, 1995, 3640, 6965, 4700, 375, 7459, 2973 , 6010, 7585, 4463, 1779, 7433, 546, 7273, 6707, 2556, 1030, 8158, 7940, 5475, 8633, 6917, 8630, 8730 , 8869, 1174, 2878, 2679, 7971, 2555, 1640, 5009, 4793, 8829, 164, 1761, 7531, 3008, 4205, 6484, 8267 , 2531, 3447, 4223, 3893, 8891, 1561, 5197, 6241, 4272, 3284, 7462, 8855, 377, 6176, 8024, 3131, 5857 , 8909, 6621, 6187, 6646, 4338, 6985, 4005, 8352, 2336, 6325, 4068, 2325, 983, 3438, 3971, 5673, 6198 , 8842, 6939, 2267, 7439, 8928, 6927, 1438, 2191, 7114, 5135, 1435, 5146, 1841, 4658, 4708, 191, 212, 4932, 2098, 1929, 2807, 6138, 3047, 5669, 1441, 1671 , 4728, 8566, 799, 4202, 3164, 8470, 8200, 8768, 5195, 3156, 8567, 1439, 2062, 635, 2518, 5705, 3473 , 8228, 8091, 7784, 190, 4475, 7551, 7094, 3302, 4851, 8672, 5352, 7464, 974, 5147, 982, 5987, 2075, 7682, 2363, 211, 8898, 5093, 4337, 8797, 7431, 8257 , 8296, 6000, 2494, 936, 3265, 4262, 1388, 2630, 8701, 1765, 888, 4036, 8551, 1845, 6303, 1427, 1351 , 6433, 4115, 8738, 1834, 586, 1043, 2131, 5207, 8281, 6489, 6221, 6681, 5750, 8722, 2731, 7555, 5112 , 5671, 6879, 7134, 4094, 3661, 4144, 7536, 5851, 6380, 3547, 5832, 1399, 6987, 1054, 6472, 2013, 7693 , 5597, 7773, 1011, 7081, 4282, 2470, 4153, 3391,
-----	---

310 3722, 1140, 364, 5413, 1430, 5013, 399, 1751, 8162, 2925, 4860, 1232, 6270, 7544, 1025, 1987, 94, 7153, 6542, 8973, 6163, 8063, 580, 1651, 1070, 8879, 2875 , 4600, 8229, 4649, 6440, 3059, 8691, 1786, 128, 5788, 1286, 2668, 8799, 985, 3649, 8895, 8107, 7196 , 6233, 8540, 8518, 3749, 5455, 197, 3914, 5218, 4950, 3859, 1114, 4019, 3615, 3988, 8726, 7992, 5602 , 7806, 1273, 5343, 7160, 7275, 8803, 438, 1778, 1465, 4711, 6712, 421, 6468, 3143, 868, 155, 794, 4979, 5053, 7258, 6913, 2920, 4500, 8659, 6560, 2820 , 2476, 8092, 6792, 6641, 3321, 734, 6779, 3468, 7173, 8127, 8343, 7597, 3751, 5740, 7176, 8950, 4546 , 3608, 1603, 1939, 6199, 7914, 5757, 388, 4565, 7297, 2309, 3965, 6262, 3125, 581, 5915, 1805, 109, 5511, 644, 8286, 883, 2682, 154, 1687, 3219, 1922, 4928, 5118, 3671, 4288, 1569, 4893, 8809, 8387, 5742 , 7370, 6704, 6932, 2742, 1574, 5528, 4285, 4618, 1062, 52, 8340, 1378, 5007, 2136, 3481, 4422, 2054, 8144, 7199, 8952, 6103, 4814, 2891, 7549, 2146, 929 , 7281, 2791, 1274, 7685, 3579, 2962, 7023, 8489, 7908, 8179, 6287, 6273, 629, 3133, 7473, 2636, 8364 , 7216, 4581, 8148, 1637, 8455, 5564, 44, 6893, 3822 , 3805, 6424, 7257, 4863, 3573, 7239, 6428, 7813, 3742, 6702, 7009, 8351, 2704, 4442, 1264, 1014, 3346 , 7608, 2735, 3376, 2931, 1341, 8052, 6890, 5391, 6569, 3203, 151, 7740, 6731, 4668, 1635, 6503, 3549 , 5080, 2978, 2155, 5991, 3149, 6226, 3731, 5846, 8256, 6351, 878, 8349, 4808, 3870, 4904, 3840, 1374 , 2907, 3305, 7664, 2497, 2769, 6013, 4709, 7984, 3993, 1578, 7487, 218, 2564, 1443, 3054, 348, 385, 1799, 2961, 3260, 8428, 8902, 5880, 2603, 6376, 501 , 1683, 3508, 956, 4836, 1946, 8375, 1741, 7029, 4683, 8026, 4615, 7342, 3913, 7490, 519, 3316, 5588 , 4773, 773, 3099, 6743, 3458, 7080, 5642, 7445, 1254, 8490, 5520, 5766, 8088, 2464, 58, 4734, 3444, 5521, 6819, 2870, 5110, 8707, 7139, 4170, 8811, 5617 , 7219, 3842, 5322, 4207, 4452, 1080, 5647, 2032, 6700, 3869, 2022, 8591, 7625, 4800, 4107, 1835, 1705 , 3641, 8759, 2515, 6315, 4216, 8402, 3003, 4718, 6361, 7720, 4027, 4235, 780, 4831, 5918, 5905, 2423 , 8554, 8737, 8576, 5512, 1160, 2743, 267, 4828,

310 6313, 4910, 3158, 1808, 1658, 1310, 1991, 1674, 6405  
, 5631, 8005, 6883, 2291, 2811, 117, 3901, 5920,  
4689, 1701, 6892, 7166, 3137, 8667, 2256, 1980, 5142  
, 7829, 3285, 3740, 2496, 621, 884, 7612, 6015, 7316  
, 5092, 7441, 4031, 801, 7588, 944, 6737, 3407, 7358  
, 764, 2608, 6858, 3607, 1794, 4722, 3789, 5976,  
5794, 5052]  
311 Label: 1  
312 generate a string that only contains the words in  
the text that are only semantically equal to missing  
trapped or found:  
313  
314 Kashmir Floods Accurate warnings issued by Flood  
Control Dept of Jammu in 2010 ignored by 'secular'  
Congress and NC http Northern California earthquake  
wipes out power to tens of thousands http http 2.9  
lakh people in 294 villages remain marooned in flood  
waters in Odisha ETTN Monsoon floods hit north  
India, 1,500 villages under water: Thousands were  
marooned in villages in UP, wh... http

```
322 6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1. Subsequent calls will reuse this data.  
323 Downloading data files: 100%|██████████| 1/1 [00:00  
<?, ?it/s]  
324 Extracting data files: 100%|██████████| 1/1 [00:00<  
00:00, 149.65it/s]  
325 Generating train split: 0 examples [00:00, ?  
examples/s]Downloading and preparing dataset csv/  
default to C:/Users/admin/.cache/huggingface/  
datasets/csv/default-e8797bce58b21435/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1...  
326 100%|██████████| 1/1 [00:00<00:00, 91.19it/s]  
327 Dataset csv downloaded and prepared to C:/Users/  
admin/.cache/huggingface/datasets/csv/default-  
e8797bce58b21435/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1. Subsequent calls will reuse this data.  
328 Downloading data files: 100%|██████████| 1/1 [00:00  
<?, ?it/s]  
329 Extracting data files: 100%|██████████| 1/1 [00:00<  
00:00, 148.23it/s]  
330 Generating train split: 0 examples [00:00, ?  
examples/s]Downloading and preparing dataset csv/  
default to C:/Users/admin/.cache/huggingface/  
datasets/csv/default-430676940ea5d79a/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1...  
331 100%|██████████| 1/1 [00:00<00:00, 100.33it/s]  
332 Dataset csv downloaded and prepared to C:/Users/  
admin/.cache/huggingface/datasets/csv/default-  
430676940ea5d79a/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1. Subsequent calls will reuse this data.  
333 Downloading data files: 100%|██████████| 1/1 [00:00  
<?, ?it/s]  
334 Downloading and preparing dataset csv/default to C:/  
Users/admin/.cache/huggingface/datasets/csv/default-  
1d6a29e35cd65be4/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
```

```
334 3bcc7ec513e1...
335 Extracting data files: 100%|██████████| 1/1 [00:00<
  00:00, 139.18it/s]
336 100%|██████████| 1/1 [00:00<00:00, 91.27it/s]
337 Dataset csv downloaded and prepared to C:/Users/
  admin/.cache/huggingface/datasets/csv/default-
  1d6a29e35cd65be4/0.0.0/
  6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
  3bcc7ec513e1. Subsequent calls will reuse this data.
338 t is : 2
339 addlen is: 1000
340 noexp_df.shape[0] last 7195
341 unexplained dataset length is:
342 7195
343 embedding begin
344 trian and val dataset length: 79200
345 diversity_sampling_openai
346 diversity_sampling_openai
347 diversity_sampling_openai
348 diversity_sampling_openai
349 diversity_sampling_openai
350 diversity_sampling_openai
351 diversity_sampling_openai
352 diversity_sampling_openai
353 720.0
354 ./embeddings/NEW_bertie_embeddings_textattack/bert-
  base-uncased-MNLI_subset_1.pt
355 Found cached dataset csv (C:/Users/admin/.cache/
  huggingface/datasets/csv/default-494a068ef8ae599d/0.
  0.0/
  6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
  3bcc7ec513e1)
356 100%|██████████| 1/1 [00:00<00:00, 184.06it/s]
357 test len 17655
358 Downloading and preparing dataset csv/default to C:/
  Users/admin/.cache/huggingface/datasets/csv/default-
  b2524976c250f856/0.0.0/
  6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
  3bcc7ec513e1...
359 Downloading data files: 100%|██████████| 1/1 [00:00<
  00:00, 1966.39it/s]
```

```
360 Extracting data files: 100%|██████████| 1/1 [00:00<
  00:00, 127.36it/s]
361 100%|██████████| 1/1 [00:00<00:00, 90.96it/s]
362 Dataset csv downloaded and prepared to C:/Users/
  admin/.cache/huggingface/datasets/csv/default-
  b2524976c250f856/0.0.0/
  6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
  3bcc7ec513e1. Subsequent calls will reuse this data.
363 preprocess_samples begin ...
364 uncertainty shape: 1530
365 ./test_embeddings/NEW_bertie_embeddings_textattack/
  bert-base-uncased-MNLI_subset_1.pt
366 classify begin
367 7257
368 0%|          | 0/20 [00:00<?, ?it/s]
  diversity_sampling_openai
369 diversity_sampling_openai
370 795it [00:11, 289.70it/s]diversity_sampling_openai
371 diversity_sampling_openai
372 810it [00:21, 289.70it/s]diversity_sampling_openai
373 diversity_sampling_openai
374 1596it [00:30, 570.76it/s]diversity_sampling_openai
375 diversity_sampling_openai
376 diversity_sampling_openai
377 1620it [00:41, 570.76it/s]diversity_sampling_openai
378 2412it [00:47, 334.35it/s]diversity_sampling_openai
379 diversity_sampling_openai
380 diversity_sampling_openai
381 2430it [01:01, 334.35it/s]diversity_sampling_openai
382 3191it [01:07, 381.78it/s]diversity_sampling_openai
383 diversity_sampling_openai
384 diversity_sampling_openai
385 3241it [01:23, 10.10it/s] diversity_sampling_openai
386 4035it [01:26, 315.94it/s]diversity_sampling_openai
387 diversity_sampling_openai
388 diversity_sampling_openai
389 4051it [01:43, 5.63it/s] diversity_sampling_openai
390 4826it [01:45, 416.82it/s]diversity_sampling_openai
391 diversity_sampling_openai
392 diversity_sampling_openai
393 diversity_sampling_openai
```

```
394 5649it [02:02, 407.88it/s]diversity_sampling_openai
395 diversity_sampling_openai
396 diversity_sampling_openai
397 diversity_sampling_openai
398 6430it [02:19, 381.16it/s]diversity_sampling_openai
399 diversity_sampling_openai
400 diversity_sampling_openai
401 6481it [02:34, 9.58it/s] diversity_sampling_openai
402 7267it [02:36, 308.44it/s]diversity_sampling_openai
403 diversity_sampling_openai
404 diversity_sampling_openai
405 diversity_sampling_openai
406 8089it [02:53, 351.17it/s]diversity_sampling_openai
407 diversity_sampling_openai
408 diversity_sampling_openai
409 diversity_sampling_openai
410 8885it [03:10, 304.41it/s]diversity_sampling_openai
411 diversity_sampling_openai
412 diversity_sampling_openai
413 8911it [03:26, 6.39it/s] diversity_sampling_openai
414 9704it [03:28, 454.34it/s]diversity_sampling_openai
415 diversity_sampling_openai
416 diversity_sampling_openai
417 9721it [03:45, 7.89it/s] diversity_sampling_openai
418 10496it [03:48, 309.53it/s]diversity_sampling_openai
419 diversity_sampling_openai
420 diversity_sampling_openai
421 10531it [04:03, 7.47it/s] diversity_sampling_openai
422 11306it [04:05, 459.69it/s]diversity_sampling_openai
423 diversity_sampling_openai
424 diversity_sampling_openai
425 diversity_sampling_openai
426 12144it [04:21, 452.73it/s]diversity_sampling_openai
427 diversity_sampling_openai
428 diversity_sampling_openai
429 12151it [04:35, 8.88it/s] diversity_sampling_openai
430 12928it [04:37, 347.77it/s]diversity_sampling_openai
431 diversity_sampling_openai
432 diversity_sampling_openai
433 12961it [04:53, 7.92it/s] diversity_sampling_openai
434 13726it [04:55, 459.18it/s]diversity_sampling_openai
```

```
435 diversity_sampling_openai
436 diversity_sampling_openai
437 diversity_sampling_openai
438 14576it [05:11, 398.18it/s]diversity_sampling_openai
439 diversity_sampling_openai
440 14580it [05:22, 398.18it/s]diversity_sampling_openai
441 diversity_sampling_openai
442 15379it [05:28, 392.33it/s]diversity_sampling_openai
443 diversity_sampling_openai
444 diversity_sampling_openai
445 15391it [05:43, 7.16it/s] diversity_sampling_openai
446 16183it [05:45, 339.01it/s]diversity_sampling_openai
447 diversity_sampling_openai
448 epoch number 19
449 diversity_sampling_openai
450 diversity_sampling_openai
451 diversity_sampling_openai
452 diversity_sampling_openai
453 16200it [06:02, 339.01it/s]global t is : 2
454 test epoch results {'accuracy': 0.5483660130718955
, 'f1_weighted': 0.5122576271107121, 'f1_macro': 0.
4106672775339226, 'f1': array([0.81 , 0.
04761905, 0.42105263, 0.20779221, 0.5389755 ,
0.04347826, 0.50171821, 0.58813559, 0.
53723404]), 'precision': array([0.7902439 , 0.2
, 0.59259259, 0.53333333, 0.51489362,
0.66666667, 0.62393162, 0.47664835, 0.
56111111]), 'recall': array([0.83076923, 0.02702703
, 0.32653061, 0.12903226, 0.56542056,
0.02247191, 0.41954023, 0.76769912, 0.
51530612])}
455
456
457
458 val epoch results {'accuracy': 0.575, 'f1_weighted'
: 0.5368155604216756, 'f1_macro': 0.
3994226680468992, 'f1': array([0.83928571, 0
. , 0.22222222, 0.2972973 , 0.57142857,
0. , 0.50359712, 0.60144928, 0.
55952381]), 'precision': array([0.81034483, 0
. , 0.5 , 0.6875 , 0.55769231,
0. , 0.59322034, 0.48115942, 0.
63513514]), 'recall': array([0.87037037, 0
. , 0.14285714, 0.18965517, 0.58585859,
```

```
461          0.          , 0.4375      , 0.80193237, 0.5
        ])}
462 16200it [06:02, 44.68it/s]
463 diversity new embedding 7195
464 diversity len embedding 7195
465 MLP_1h(
466     (l1): Linear(in_features=33, out_features=100,
467      bias=True)
467     (l2): Linear(in_features=100, out_features=9, bias
468      =True)
469 [1120, 6335, 1440, 953, 797, 6098, 3173, 4099, 3535
469 , 5324, 3926, 2798, 4790, 6350, 5962, 5294, 6990,
469 1465, 6256, 6841, 4578, 1050, 3892, 7165, 5600, 3569
469 , 1414, 6304, 1973, 4629, 6834, 1193, 1199, 5187,
469 4676, 4582, 2570, 3709, 4928, 6172, 2524, 5060, 7032
469 , 3611, 5881, 1658, 319, 5184, 5364, 1472, 4028,
469 5477, 2649, 5641, 2588, 1079, 6827, 2004, 4874, 4878
469 , 2976, 4359, 3185, 6322, 4445, 6048, 5307, 6583,
469 7146, 295, 6236, 4191, 6671, 4675, 936, 503, 3427,
469 788, 2652, 4737, 2931, 6163, 159, 2117, 5810, 5196,
469 1583, 3711, 5016, 528, 4650, 6456, 2836, 2317, 701,
469 2786, 5711, 2774, 3024, 7077, 2782, 3274, 2140, 1528
469 , 1408, 1800, 3327, 5939, 3156, 1763, 3249, 564,
469 7002, 2354, 706, 4929, 6580, 1101, 4751, 2062, 2053
469 , 7033, 5348, 511, 1180, 2154, 1443, 4635, 74, 746,
469 1356, 6170, 4973, 1104, 1492, 6353, 5926, 1288, 5776
469 , 3255, 5506, 3373, 3163, 26, 5859, 2162, 3104, 3257
469 , 5760, 351, 3653, 669, 3872, 5345, 7113, 5615, 2139
469 , 1100, 3404, 1722, 4836, 3401, 60, 5310, 2662, 2292
469 , 4769, 3731, 5232, 817, 3179, 1271, 5664, 1426,
469 6758, 4364, 6249, 3950, 4153, 1367, 3955, 1945, 4390
469 , 6685, 2844, 4091, 2658, 5538, 1578, 4560, 37, 6337
469 , 6547, 1637, 4733, 6168, 507, 4520, 2631, 1085,
469 3480, 6417, 6199, 4365, 2289, 220, 621, 1231, 2186,
469 5759, 2978, 1636, 6678, 5243, 3080, 3699, 2699, 96,
469 6177, 6122, 544, 2504, 2728, 3074, 1885, 1240, 5709
469 , 442, 3963, 1491, 50, 5640, 1312, 853, 2867, 4182,
469 4109, 6196, 6239, 2824, 4300, 6821, 2636, 1219, 6918
469 , 3507, 455, 1017, 330, 1566, 4824, 5125, 15, 4043,
469 2272, 5522, 3304, 1921, 4517, 90, 5067, 1643, 5142,
```

469 6786, 3351, 4040, 5299, 4672, 210, 3071, 1907, 3131  
, 6750, 5716, 1027, 5757, 1313, 3501, 4069, 4587,  
2415, 5877, 6718, 2434, 6062, 5887, 3672, 6057, 5172  
, 6479, 4952, 3475, 3882, 2994, 4940, 1771, 1876,  
4074, 95, 3034, 6221, 4544, 5799, 6975, 2445, 5393,  
3443, 2977, 6860, 3645, 2565, 1677, 5341, 4081, 1355  
, 4362, 6866, 4860, 1815, 6603, 4994, 4016, 355,  
2159, 2512, 1245, 3770, 1520, 6762, 6795, 2071, 2137  
, 4586, 3839, 4564, 6566, 894, 3548, 4287, 5751,  
4557, 1484, 1212, 1616, 477, 3625, 2286, 7020, 291,  
6068, 7036, 5566, 6816, 3876, 4491, 5002, 4222, 3747  
, 5097, 2606, 5567, 4329, 3806, 2871, 223, 6441, 845  
, 1672, 6426, 883, 1235, 1341, 4010, 5034, 3422,  
4562, 856, 7072, 6305, 1667, 5862, 6969, 4503, 1328  
, 979, 394, 5314, 6161, 3364, 5113, 6949, 6399, 594  
, 3572, 5042, 7031, 6132, 1797, 6731, 4367, 3315,  
4314, 269, 1241, 5958, 4171, 5417, 3453, 4710, 4084  
, 135, 1901, 506, 75, 768, 1767, 4707, 2166, 5978,  
2035, 3980, 3736, 2950, 5086, 5107, 4094, 3002, 1139  
, 181, 6797, 1351, 6397, 3740, 3227, 5446, 2107, 806  
, 5552, 3094, 5621, 2284, 4971, 4159, 4351, 6018,  
2099, 1068, 17, 7157, 2248, 3706, 2280, 6008, 4854,  
804, 1535, 6532, 2593, 5338, 1146, 5988, 4342, 5409  
, 6389, 7123, 2492, 2668, 5660, 4609, 2518, 4183,  
6446, 429, 2981, 6325, 1813, 1717, 1673, 5106, 3734  
, 3198, 1466, 364, 4216, 4763, 5249, 1407, 2736,  
6716, 323, 3243, 2462, 6127, 5423, 192, 2323, 245,  
805, 5896, 1444, 202, 4221, 3177, 1858, 4166, 886,  
2739, 2908, 2897, 6688, 5226, 5181, 2731, 1381, 570  
, 6103, 2999, 4484, 2919, 2357, 4492, 7017, 1267,  
1183, 2726, 4932, 7096, 7097, 5280, 3493, 1682, 1888  
, 1521, 132, 1298, 1727, 8, 4443, 673, 3712, 4853,  
6836, 6945, 3566, 6481, 4161, 4931, 3450, 3510, 6719  
, 2812, 1962, 5911, 3647, 3869, 5833, 4425, 3598,  
4251, 1811, 6592, 3814, 562, 4680, 3429, 1010, 2165  
, 5658, 2961, 5350, 2572, 5156, 2586, 1721, 3944,  
2856, 2433, 4036, 172, 4024, 4206, 2090, 6878, 2609  
, 6571, 1738, 4293, 374, 3788, 4261, 4379, 1806,  
2895, 2358, 1201, 4600, 6814, 3190, 209, 965, 3506,  
5724, 3651, 2017, 6572, 904, 5404, 1224, 115, 929,  
3675, 4529, 5870, 1251, 4596, 1186, 4035, 3722, 2399

469 , 1141, 1232, 201, 560, 5633, 6513, 2510, 378, 803, 5102, 4735, 531, 5837, 5608, 360, 1284, 2468, 6737, 659, 5873, 5158, 6887, 774, 5612, 3659, 1635, 2566, 6973, 5871, 1070, 3660, 3084, 5269, 7177, 5695, 1105 , 908, 4196, 4585, 5137, 2963, 5065, 1881, 6553, 4452, 6087, 4912, 3870, 5900, 6626, 4721, 472, 7099 , 2575, 6091, 6537, 1459, 5955, 1323, 6736, 1359, 4231, 2880, 1494, 4855, 1683, 2659, 5949, 5588, 2892 , 3457, 5117, 540, 4478, 2046, 4048, 7073, 4797, 5562, 111, 203, 1563, 3903, 6015, 4659, 1369, 2491, 2529, 4189, 6214, 7194, 5467, 6627, 7185, 3863, 4766 , 5910, 7088, 5628, 1949, 2291, 4713, 5880, 609, 5360, 1920, 1739, 3117, 5136, 5601, 5892, 1669, 1306 , 6190, 5842, 2018, 1214, 1136, 3495, 1626, 5359, 6796, 5735, 3345, 6617, 3584, 945, 2235, 3417, 5238 , 3564, 4825, 1774, 2806, 247, 418, 6071, 2240, 5942 , 5440, 1526, 6001, 5675, 1715, 4291, 2794, 6486, 2370, 6680, 4187, 3348, 4958, 6266, 3057, 2579, 468 , 3750, 6200, 1986, 2143, 4512, 339, 3272, 4946, 6916, 4806, 2372, 1996, 5491, 3306, 6121, 2223, 5544 , 1608, 3902, 2653, 1840, 6225, 7188, 3737, 1567, 1590, 7052, 393, 1519, 1237, 1780, 6904, 4726, 2245 , 5919, 1979, 5301, 625, 3019, 5809, 6333, 6152, 5197, 891, 3755, 68, 5277, 3601, 5296, 3891, 3767, 4872, 6061, 5072, 569, 212, 5358, 3431, 5690, 5462, 4453, 1188, 1966, 2296, 5831, 6847, 6898, 6811, 4457 , 5311, 6899, 6065, 802, 4013, 1971, 3049, 6558, 5952, 551, 5425, 2818, 3757, 2545, 1198, 1548, 7118 , 6260, 4008, 2014, 4285, 1257, 2405, 500, 7038, 2198, 5902, 3310, 4739, 643, 3631, 7108, 6609, 948, 2709, 1644, 6641, 980, 4395, 5940, 6701, 1483, 6113 , 1869, 1138, 2274, 1708, 160, 2822, 5368, 1387, 6843, 2229, 185, 3531, 6519, 1112, 5700, 7034, 6782 , 3754, 5950, 2016, 807, 2373, 5229, 1057, 1594, 5836, 5584, 5085, 3194, 4079, 7000, 793, 5285, 3929 , 611, 2064, 2130, 745, 2461, 4325, 4636, 213, 4844 , 1585, 3692, 2457, 5354, 4799, 521, 1067, 964, 4199 , 3707, 6395, 6735, 2563, 4945, 3323, 3250, 5792, 5037, 1604, 2738, 6630, 2138, 1353, 942, 4840, 3253 , 4755, 4942, 5846, 799, 3252, 1065, 6658, 6316, 533 , 6354, 3013, 2800, 4032, 3799, 4353, 5415, 4219,

469 4082, 722, 6687, 2942, 4156, 3260, 5041, 5239, 2992  
, 5590, 4734, 1882, 5497, 1556, 5931, 4322, 4393,  
5157, 2747, 2255, 6060, 2371, 2741, 3843, 4683, 4563  
, 2980, 2706, 5519, 6506, 6852, 2199, 4882, 157, 840  
, 2056, 1296, 6531, 4117, 1308, 2167, 7091, 5532,  
2383, 2832, 3093, 5441, 5135, 6023, 5696, 3796, 3619  
, 4273, 5082, 307, 1842, 415, 4785, 263, 4569, 4627  
, 519, 717, 4253, 833, 262, 1933, 2900, 280, 1618,  
4832, 3801, 4916, 6197, 2243, 2646, 3689, 4679, 843  
, 6576, 1258, 1265, 689, 7139, 2721, 4496, 5669,  
4381, 5912, 5539, 7074, 3030, 1125, 231, 6157, 2677  
, 5505, 4757, 574, 2227, 6382, 5857, 6499, 5813,  
6004, 5619, 588, 6255, 2493, 1133, 237, 4374, 770,  
5363, 6409, 5864, 1875, 6028, 6695, 4540, 5331, 5798  
, 6655, 1294, 6115, 1041, 4704, 2439, 4620, 3784,  
6647, 915, 4745, 2790, 2113, 2257, 4652, 1837, 3465  
, 4076, 6024, 6689, 3143, 47, 1941, 10, 1168, 2511,  
6470, 3288, 2848, 101, 938, 4185, 4779, 1315, 5727,  
4896, 2729, 5749, 5216, 6429, 950, 849, 6414, 433,  
1073, 3456, 4746, 2777, 1726, 1015, 3088, 5835, 2076  
, 5909, 3866, 4438, 2091, 19, 3988, 5802, 5174, 4447  
, 5513, 4164, 56, 1044, 2845, 3986, 2412, 3142, 4236  
, 1230, 2864, 5329, 1542, 5019, 3363, 5548, 6941,  
6633, 5129, 4674, 3798, 98, 4317, 7059, 6274, 3698,  
6261, 1792, 380, 3780, 4174, 5453, 4025, 6854, 3494  
, 6450, 6039, 3236, 3562, 1038, 2256, 5895, 1263,  
5787, 1000, 3115, 6517, 6303, 693, 420, 118, 5845,  
4724, 3761, 971, 3803, 1980, 2124, 4993, 1476, 5391  
, 3331, 5681, 1853, 4701, 2151, 5551, 5647, 1712,  
5795, 3420, 149, 5346, 703, 5308, 5662, 2630, 5089,  
4269, 2044, 6828, 2998, 4862, 3499, 3481, 1170, 1211  
, 2714, 170, 5406, 1911, 1221, 2948, 3648, 1699, 24  
, 6909, 2058, 1270, 6386, 6494, 4179, 3368, 5917,  
3220, 5046, 927, 7035, 5033, 7191, 4549, 6602, 4140  
, 6994, 3040, 1121, 658, 3538, 4203, 4405, 2428,  
3283, 1266, 5096, 6058, 7005, 4200, 6405, 5483, 6176  
, 5055, 224, 2049, 1005, 2725, 404, 4850, 4262, 6026  
, 28, 3524, 3468, 2862, 2828, 4791, 423, 5408, 3047  
, 5508, 494, 1782, 2710, 2803, 6625, 6637, 1631,  
1956, 5854, 4615, 152, 4077, 4347, 4786, 2250, 4917  
, 4869, 6706, 5304, 5656, 5764, 5267, 4654, 4054,

469 7046, 6431, 3553, 4918, 6408, 579, 4802, 3899, 3447  
, 313, 7121, 3418, 127, 5790, 3438, 3367, 1054, 6331  
, 5144, 256, 649, 6524, 3896, 400, 3831, 5316, 2419  
, 4436, 6739, 615, 6657, 1544, 5297, 2142, 755, 6615  
, 2078, 3855, 7004, 2464, 6996, 3775, 3010, 1588,  
4416, 663, 7054, 1944, 1524, 7140, 961, 6970, 756,  
3676, 973, 4313, 4518, 406, 4920, 1915, 1167, 5171,  
584, 2837, 6709, 1749, 6585, 2263, 1202, 7134, 2037  
, 2041, 754, 4534, 752, 4042, 2131, 6722, 5576, 6560  
, 178, 4248, 4773, 1538, 2751, 512, 5432, 2195, 4688  
, 6971, 6276, 1034, 968, 1043, 1144, 4657, 6268,  
3840, 3370, 5211, 3372, 4471, 1541, 880, 5300, 3993  
, 6679, 3301, 4964]  
470 Label: 2  
471 generate a string that only contains the words in  
the text that are only semantically equal to  
displaced and evacuations:  
472  
473 Pakistan Floods affect more than one million http  
http Residents, Tourists Sent to Shelters as  
Hurricane Odile Makes Landfall - ABC News: ABC  
NewsResidents, Tourists... http Hagupit Oxfam rapid  
assessment teams to survey immediate needs of  
evacuees http Philippines http Communities mobilize  
for relief as Super Typhoon Hagupit roars toward  
Philippines http Ruby P H Chile - Residents of  
Iquique spend the night on the street after a second  
major quake hit the region in as many days http.  
474  
475  
476 AI Explanation: "Residents, Tourists evacuations  
Shelters Hurricane Odile makes rapid assessment  
teams survey immediate needs of evacuees Communities  
477 Downloading and preparing dataset csv/default to C:/  
Users/admin/.cache/huggingface/datasets/csv/default-  
76a32e9fb1f2c4ea/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1...  
478 Downloading data files: 100%|██████████| 1/1 [00:00  
<?, ?it/s]  
479 Extracting data files: 100%|██████████| 1/1 [00:00<

```
479 00:00, 143.74it/s]
480 100%|██████████| 1/1 [00:00<00:00, 91.22it/s]
481 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
76a32e9fb1f2c4ea/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1. Subsequent calls will reuse this data.
482 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
483 Downloading and preparing dataset csv/default to C:/
Users/admin/.cache/huggingface/datasets/csv/default-
4dc88927f1b58eb0/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1...
484 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 113.06it/s]
485 100%|██████████| 1/1 [00:00<00:00, 59.22it/s]
486 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
4dc88927f1b58eb0/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1. Subsequent calls will reuse this data.
487 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
488 Downloading and preparing dataset csv/default to C:/
Users/admin/.cache/huggingface/datasets/csv/default-
dc3e80cfbc1b8920/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1...
489 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 38.59it/s]
490 100%|██████████| 1/1 [00:00<00:00, 71.65it/s]
491 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
dc3e80cfbc1b8920/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1. Subsequent calls will reuse this data.
492 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
493 Downloading and preparing dataset csv/default to C:/
Users/admin/.cache/huggingface/datasets/csv/default-
```

```
493 7ff6db531f575ecf/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
    3bcc7ec513e1...
494 Extracting data files: 100%|██████████| 1/1 [00:00<
    00:00, 110.84it/s]
495 100%|██████████| 1/1 [00:00<00:00, 91.21it/s]
496 Dataset csv downloaded and prepared to C:/Users/
    admin/.cache/huggingface/datasets/csv/default-
    7ff6db531f575ecf/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
    3bcc7ec513e1. Subsequent calls will reuse this data.
497 t is : 3
498 addlen is: 1500
499 noexp_df.shape[0] last 5756
500 unexplained dataset length is:
501 5756
502 embedding begin
503 trian and val dataset length: 103680
504 diversity_sampling_openai
505 diversity_sampling_openai
506 diversity_sampling_openai
507 diversity_sampling_openai
508 diversity_sampling_openai
509 diversity_sampling_openai
510 diversity_sampling_openai
511 diversity_sampling_openai
512 864.0
513 ./embeddings/NEW_bertie_embeddings_textattack/bert-
    base-uncased-MNLI_subset_1.pt
514 Found cached dataset csv (C:/Users/admin/.cache/
    huggingface/datasets/csv/default-494a068ef8ae599d/0.
    0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
    3bcc7ec513e1)
515 100%|██████████| 1/1 [00:00<00:00, 59.52it/s]
516 test len 19260
517 Downloading and preparing dataset csv/default to C:/
    Users/admin/.cache/huggingface/datasets/csv/default-
    8c4c6c705fbc7426/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
    3bcc7ec513e1...
```

```
518 Downloading data files: 100%|██████████| 1/1 [00:00
    <?, ?it/s]
519 Extracting data files: 100%|██████████| 1/1 [00:00<
    00:00, 127.35it/s]
520 100%|██████████| 1/1 [00:00<00:00, 71.70it/s]
521 Dataset csv downloaded and prepared to C:/Users/
    admin/.cache/huggingface/datasets/csv/default-
    8c4c6c705fbc7426/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
    3bcc7ec513e1. Subsequent calls will reuse this data.
522 preprocess_samples begin ...
523 uncertainty shape: 1530
524 ./test_embeddings/NEW_bertie_embeddings_textattack/
    bert-base-uncased-MNLI_subset_1.pt
525 classify begin
526 8696
527 0%|          | 0/20 [00:00<?, ?it/s]
    diversity_sampling_openai
528 diversity_sampling_openai
529 948it [00:11, 427.21it/s]diversity_sampling_openai
530 diversity_sampling_openai
531 diversity_sampling_openai
532 diversity_sampling_openai
533 1919it [00:28, 449.33it/s]diversity_sampling_openai
534 diversity_sampling_openai
535 diversity_sampling_openai
536 1945it [00:43, 9.03it/s] diversity_sampling_openai
537 2912it [00:45, 439.50it/s]diversity_sampling_openai
538 diversity_sampling_openai
539 diversity_sampling_openai
540 2917it [01:00, 7.43it/s] diversity_sampling_openai
541 3863it [01:02, 425.64it/s]diversity_sampling_openai
542 diversity_sampling_openai
543 diversity_sampling_openai
544 diversity_sampling_openai
545 4840it [01:19, 388.25it/s]diversity_sampling_openai
546 diversity_sampling_openai
547 4860it [01:30, 388.25it/s]diversity_sampling_openai
548 diversity_sampling_openai
549 5825it [01:36, 431.51it/s]diversity_sampling_openai
550 diversity_sampling_openai
```

```
551 diversity_sampling_openai
552 5833it [01:51,  7.73it/s] diversity_sampling_openai
553 6773it [01:54, 305.78it/s] diversity_sampling_openai
554 diversity_sampling_openai
555 diversity_sampling_openai
556 diversity_sampling_openai
557 7767it [02:11, 434.63it/s] diversity_sampling_openai
558 diversity_sampling_openai
559 diversity_sampling_openai
560 diversity_sampling_openai
561 8735it [02:28, 505.92it/s] diversity_sampling_openai
562 diversity_sampling_openai
563 diversity_sampling_openai
564 8749it [02:42,  9.75it/s] diversity_sampling_openai
565 9703it [02:44, 561.04it/s] diversity_sampling_openai
566 diversity_sampling_openai
567 diversity_sampling_openai
568 diversity_sampling_openai
569 10674it [03:02, 444.44it/s] diversity_sampling_openai
570 diversity_sampling_openai
571 diversity_sampling_openai
572 diversity_sampling_openai
573 11662it [03:19, 424.31it/s] diversity_sampling_openai
574 diversity_sampling_openai
575 diversity_sampling_openai
576 11665it [03:33,  7.52it/s] diversity_sampling_openai
577 12615it [03:36, 435.07it/s] diversity_sampling_openai
578 diversity_sampling_openai
579 diversity_sampling_openai
580 12637it [03:50,  8.64it/s] diversity_sampling_openai
581 13600it [03:54, 273.04it/s] diversity_sampling_openai
582 diversity_sampling_openai
583 diversity_sampling_openai
584 diversity_sampling_openai
585 14563it [04:11, 437.39it/s] diversity_sampling_openai
586 diversity_sampling_openai
587 diversity_sampling_openai
588 diversity_sampling_openai
589 15534it [04:28, 409.54it/s] diversity_sampling_openai
590 diversity_sampling_openai
591 diversity_sampling_openai
```

```
592 15553it [04:43, 8.04it/s] diversity_sampling_openai
593 16496it [04:46, 398.35it/s] diversity_sampling_openai
594 diversity_sampling_openai
595 diversity_sampling_openai
596 16525it [05:00, 8.53it/s] diversity_sampling_openai
597 17494it [05:03, 431.95it/s] diversity_sampling_openai
598 diversity_sampling_openai
599 diversity_sampling_openai
600 diversity_sampling_openai
601 18453it [05:20, 312.36it/s] diversity_sampling_openai
602 diversity_sampling_openai
603 18468it [05:30, 312.36it/s] diversity_sampling_openai
604 diversity_sampling_openai
605 19418it [05:37, 537.65it/s] diversity_sampling_openai
606 diversity_sampling_openai
607 epoch number 19
608 diversity_sampling_openai
609 diversity_sampling_openai
610 19440it [05:50, 537.65it/s] diversity_sampling_openai
611 diversity_sampling_openai
612 global t is : 3
613 test epoch results {'accuracy': 0.569281045751634, 'f1_weighted': 0.5413020629184628, 'f1_macro': 0.46436880884427545, 'f1': array([0.85148515, 0.24, 0.52173913, 0.24539877, 0.56444444, 0.08333333, 0.52830189, 0.59687787, 0.54773869]), 'precision': array([0.82296651, 0.46153846, 0.55813953, 0.51282051, 0.53813559, 0.57142857, 0.58333333, 0.51020408, 0.53960396]), 'recall': array([0.88205128, 0.16216216, 0.48979592, 0.16129032, 0.59345794, 0.04494382, 0.48275862, 0.71902655, 0.55612245])}
614 val epoch results {'accuracy': 0.5590277777777778, 'f1_weighted': 0.5304509760620215, 'f1_macro': 0.4297748304345315, 'f1': array([0.86567164, 0.15384615, 0.36842105, 0.22, 0.45783133, 0.14814815, 0.50847458, 0.61101836, 0.53456221]), 'precision': array([0.82269504, 0.28571429, 0.53846154, 0.42307692, 0.43846154, 0.57142857, 0.5625, 0.51988636, 0.]}
```

```
619 53703704]), 'recall': array([0.91338583, 0.10526316
   , 0.28      , 0.14864865, 0.4789916 ,
620      0.08510638, 0.46391753, 0.74089069, 0.
   53211009])}
621 19440it [05:56, 54.53it/s]
622 diversity new embedding 5756
623 diversity len embedding 5756
624 MLP_1h(
625     (l1): Linear(in_features=36, out_features=100,
   bias=True)
626     (l2): Linear(in_features=100, out_features=9, bias
   =True)
627 )
628 [3198, 5284, 1995, 1263, 5618, 2932, 3861, 2291,
   3076, 4266, 614, 1305, 2935, 779, 4020, 5680, 120,
   4413, 1357, 1977, 48, 2639, 5019, 5068, 3608, 5712,
   2358, 1351, 4233, 5444, 4388, 2, 1174, 2163, 1660,
   2341, 5487, 3093, 406, 5024, 874, 855, 3363, 4583,
   2113, 4255, 4087, 3991, 2595, 3445, 1271, 1883, 4532
   , 3968, 3059, 2627, 5028, 2579, 3791, 1032, 3587,
   4303, 549, 3684, 1258, 503, 3718, 5297, 2709, 5258,
   174, 950, 293, 5123, 2384, 2752, 1683, 5011, 2287,
   648, 5500, 4437, 2707, 3957, 3413, 5112, 1236, 4317
   , 5022, 4242, 4085, 1415, 881, 3918, 3821, 4174,
   1207, 3710, 5281, 4027, 1933, 807, 3914, 5519, 4260
   , 194, 5239, 1702, 5362, 5375, 580, 3931, 5598, 1896
   , 1863, 3439, 4348, 774, 5398, 4782, 2603, 4639,
   1663, 3471, 1061, 3505, 4808, 5096, 3912, 2355, 4454
   , 271, 5360, 4098, 4426, 1185, 4682, 3123, 1383, 620
   , 5047, 4083, 1340, 709, 3992, 3356, 4665, 4058, 58
   , 4059, 5526, 1021, 195, 3712, 4602, 62, 1725, 4910
   , 2489, 1737, 1356, 2182, 1937, 1339, 808, 3750,
   5450, 1528, 5651, 5415, 4019, 1337, 488, 103, 1269,
   3355, 4412, 4381, 111, 1100, 1734, 4230, 4145, 1245
   , 5498, 3349, 3593, 5373, 5652, 3772, 5533, 3211,
   3293, 5669, 3002, 5614, 1543, 899, 3614, 720, 3176,
   4049, 4247, 2084, 5182, 4887, 3888, 5746, 5235, 5382
   , 1155, 2929, 3069, 1785, 2866, 1202, 449, 887, 3581
   , 285, 1126, 1592, 1789, 3663, 546, 2192, 4327, 1011
   , 5139, 5214, 2733, 1900, 3530, 1280, 4369, 3250,
   2331, 4318, 3021, 595, 3339, 5548, 5700, 2703, 2380
```

628 , 2608, 2819, 1919, 2868, 3465, 1498, 3624, 38, 2592  
, 3517, 756, 4881, 5225, 5299, 803, 4036, 2425, 377  
, 2534, 5744, 2185, 3188, 1067, 2135, 312, 4592, 464  
, 4194, 839, 2190, 4647, 3245, 4871, 1373, 2574,  
2572, 403, 5532, 3873, 4967, 1694, 991, 5317, 3969,  
4669, 5496, 229, 4494, 4405, 1151, 4581, 2323, 31,  
4956, 696, 971, 5414, 898, 2604, 5057, 455, 3467,  
2165, 5729, 3295, 1318, 5257, 2681, 5290, 4919, 5673  
, 1986, 2237, 4780, 2093, 3091, 3222, 1819, 2677,  
724, 3218, 2149, 77, 4146, 5285, 5094, 2006, 2984,  
4229, 4120, 4763, 3016, 1936, 490, 5667, 1544, 5295  
, 278, 2688, 5411, 506, 2282, 2663, 3809, 3477, 4961  
, 2144, 4055, 815, 247, 4091, 4715, 1384, 127, 1886  
, 3757, 2065, 3060, 3753, 2004, 3620, 2064, 2221,  
4721, 4393, 3337, 5111, 5274, 3984, 2669, 364, 1331  
, 5461, 3935, 1859, 684, 85, 1419, 4894, 5528, 1189  
, 708, 4909, 1194, 2656, 4010, 5439, 1405, 758, 4251  
, 2036, 3832, 1554, 1779, 4869, 3773, 1724, 726,  
3520, 3862, 613, 3051, 3725, 3351, 3023, 3206, 4556  
, 2720, 5278, 2831, 2933, 4933, 2527, 4613, 451,  
2177, 2128, 693, 5638, 2082, 3184, 35, 166, 4367,  
3266, 2120, 2941, 3589, 5706, 2755, 955, 1060, 4245  
, 3396, 834, 1988, 1606, 5170, 4419, 3500, 2451,  
2784, 2245, 5076, 900, 296, 440, 5172, 585, 2737,  
5541, 5395, 1670, 1719, 3547, 2094, 1656, 2770, 2632  
, 1532, 2025, 1680, 1489, 5302, 4991, 201, 3134,  
4332, 5359, 3032, 4440, 2563, 383, 2822, 4771, 985,  
1766, 3595, 566, 2757, 5427, 4084, 1565, 3191, 1270  
, 5241, 4485, 2337, 3592, 5724, 4495, 4188, 4978,  
5477, 2462, 155, 1484, 138, 1394, 3437, 4254, 820,  
545, 1423, 169, 4577, 5692, 5084, 1984, 3504, 5425,  
4323, 2965, 4979, 657, 1217, 3027, 1509, 538, 4857,  
987, 866, 1591, 413, 4632, 1685, 937, 2693, 3424,  
4981, 2188, 4116, 3782, 4673, 4325, 5513, 5479, 5109  
, 380, 2756, 2100, 2841, 3510, 2512, 5169, 2321,  
2200, 2676, 2252, 5592, 198, 3902, 1234, 2944, 4539  
, 4158, 645, 428, 3851, 1186, 1529, 1121, 1726, 4197  
, 4232, 5054, 1268, 2183, 493, 3308, 4018, 310, 5180  
, 4160, 4925, 2211, 891, 2046, 5021, 3479, 915, 1490  
, 942, 1364, 2257, 5639, 2124, 3707, 4178, 5604,  
5244, 468, 2057, 498, 1589, 390, 1376, 4657, 5696,

628 4110, 617, 1056, 2423, 4274, 3906, 4774, 2360, 3045  
, 5722, 4870, 393, 351, 2486, 1469, 5424, 2453, 5594  
, 1721, 4938, 3641, 515, 433, 2768, 718, 1502, 252,  
2083, 831, 653, 4565, 2329, 2104, 4314, 191, 5135,  
24, 2735, 2705, 5131, 1148, 4154, 3780, 5191, 4446,  
3549, 4813, 2810, 5624, 1560, 3287, 1306, 2171, 3473  
, 3353, 1620, 4731, 3018, 2588, 3289, 4529, 3025,  
802, 555, 308, 5013, 3626, 2943, 1510, 1774, 1445,  
2283, 5167, 3220, 2447, 1916, 3268, 4661, 3865, 853  
, 4476, 1363, 936, 4648, 3111, 2619, 5159, 2054, 21  
, 5665, 381, 4542, 1144, 4103, 5467, 3261, 4835,  
3642, 2633, 3553, 3181, 214, 5118, 3428, 4915, 1932  
, 3203, 5156, 5291, 1892, 5247, 3808, 1967, 1885,  
1760, 3147, 1943, 3761, 999, 1587, 4150, 4627, 5062  
, 2658, 2721, 567, 5715, 5530, 1644, 4179, 2526,  
2637, 1624, 461, 5447, 16, 492, 800, 1042, 786, 3868  
, 3943, 4519, 5002, 4200, 2023, 5571, 1898, 1471,  
509, 3214, 2811, 1928, 2898, 3880, 4843, 5176, 2650  
, 2320, 4265, 4460, 3468, 405, 1102, 2308, 1093,  
3512, 5542, 3673, 220, 4275, 1486, 3230, 2973, 368,  
2520, 3118, 1290, 5029, 3301, 3109, 982, 1535, 3682  
, 3242, 3185, 5426, 4359, 2953, 780, 49, 5748, 1378  
, 2394, 2444, 994, 2711, 2164, 2597, 5544, 5141,  
1156, 411, 2881, 1623, 3407, 3327, 255, 863, 1775,  
5685, 4596, 742, 2491, 4503, 2856, 145, 1041, 161,  
1583, 1550, 2769, 4754, 1152, 4559, 4726, 221, 3140  
, 2739, 1499, 4173, 2013, 1808, 1770, 319, 2505,  
1696, 3872, 5335, 4380, 844, 1374, 1028, 171, 446,  
65, 5466, 4946, 3469, 3843, 2746, 2805, 5260, 3151,  
2151, 1991, 3280, 679, 4387, 272, 2586, 1079, 1353,  
5572, 3562, 2940, 1613, 2016, 1822, 5537, 3404, 540  
, 4528, 5044, 732, 2629, 1117, 3298, 5085, 4433,  
4680, 4793, 520, 1861, 759, 3824, 2403, 1798, 331,  
5030, 1380, 4447, 4667, 553, 1427, 4211, 1066, 5372  
, 1288, 4692, 4138, 1178, 3617, 4552, 2896, 952, 681  
, 832, 4795, 2121, 1649, 2906, 5416, 3756, 1980, 840  
, 2685, 4768, 4375, 5617, 1635, 1888, 1843, 2029,  
1261, 3200, 3818, 1140, 4435, 281, 5145, 3348, 5556  
, 2095, 3677, 3249, 3270, 4165, 4077, 2132, 231, 276  
, 1927, 4872, 5516, 3022, 1199, 3386, 2448, 4945,  
4169, 3064, 3566, 4421, 2371, 130, 3048, 5714, 1466

628 , 473, 2027, 1237, 5106, 1717, 3911, 2683, 262, 3521  
, 1118, 1034, 2837, 3717, 629, 3847, 5711, 4299, 668  
, 3994, 356, 2722, 4372, 3933, 4829, 5006, 608, 663  
, 1062, 1786, 3260, 2267, 230, 2948, 3303, 2538,  
3503, 2576, 3793, 2960, 634, 5162, 2947, 4858, 1244  
, 4614, 1518, 5688, 2812, 1615, 3810, 703, 812, 5505  
, 630, 1687, 1473, 2456, 3660, 1520, 423, 4564, 1094  
, 136, 2007, 3236, 3601, 2593, 2798, 2350, 4966,  
1012, 3959, 1071, 741, 4264, 4465, 2140, 2136, 2229  
, 4700, 1051, 2386, 119, 4714, 5245, 5300, 2776,  
4710, 2585, 1664, 561, 3668, 4341, 1917, 2647, 1010  
, 3584, 1006, 1050, 4543, 1951, 469, 1671, 969, 3425  
, 5682, 1826, 1575, 2952, 1901, 3901, 5298, 1441,  
3788, 1193, 2890, 1081, 2666, 5601, 1551, 4221, 3916  
, 3161, 4093, 882, 943, 764, 3243, 3187, 3875, 1223  
, 5034, 3475, 5459, 519, 3885, 3258, 3787, 868, 922  
, 4335, 1328, 3695, 2264, 5197, 3744, 107, 2541,  
3732, 4486, 4152, 4467, 3580, 4355, 5065, 2236, 1002  
, 363, 3609, 5420, 761, 5702, 850, 1195, 329, 512,  
282, 1955, 1913, 1877, 314, 2005, 1275, 5293, 3174,  
1941, 1940, 1727, 227, 4122, 2362, 298, 5138, 1088,  
845, 5200, 3300, 1078]  
629 Label: 3  
630 generate a string that only contains the words in  
the text that are only semantically equal to  
infrastructure and utilities damage:  
631  
632 A Path of Hurricane Odile's Destruction, Seen  
Through Instagram: Hurricane Odile came ashore... http  
Give Aid to Victims of Disastrous Cyclone http  
Fortunately, Gallery 1870 in Yountville, Ca only  
incurred minor damage during Sunday's Napa Valley  
earthquake,... http (+1 849-846-0679) Scores  
Killed in Flooding in Nepal and India: More than 11,  
000 homes have been damaged and r... (( Papiiz\_  
Chulò )) My husband snapped a pic of damage through  
a window. So sad to see a beautiful suffer so much  
damage Odile ht....  
633  
634  
635 AI Explanation: "Hurricane Odile destruction, minor

```
635 damage, damaged, flooding, damaged, suffer damage,  
Odile"  
636 Downloading and preparing dataset csv/default to C:/  
Users/admin/.cache/huggingface/datasets/csv/default-  
17c24a36f7fedbb8/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1...  
637 Downloading data files: 100%|██████████| 1/1 [00:00<  
00:00, 1003.42it/s]  
638 Extracting data files: 100%|██████████| 1/1 [00:00<  
00:00, 143.33it/s]  
639 100%|██████████| 1/1 [00:00<00:00, 77.18it/s]  
640 Dataset csv downloaded and prepared to C:/Users/  
admin/.cache/huggingface/datasets/csv/default-  
17c24a36f7fedbb8/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1. Subsequent calls will reuse this data.  
641 Downloading data files: 100%|██████████| 1/1 [00:00<  
00:00, 983.42it/s]  
642 Downloading and preparing dataset csv/default to C:/  
Users/admin/.cache/huggingface/datasets/csv/default-  
bf4204d14c4cbf5e/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1...  
643 Extracting data files: 100%|██████████| 1/1 [00:00<  
00:00, 142.91it/s]  
644 0%|          | 0/1 [00:00<?, ?it/s]Dataset csv  
downloaded and prepared to C:/Users/admin/.cache/  
huggingface/datasets/csv/default-bf4204d14c4cbf5e/0.  
0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1. Subsequent calls will reuse this data.  
645 100%|██████████| 1/1 [00:00<00:00, 62.00it/s]  
646 Downloading data files: 100%|██████████| 1/1 [00:00  
<?, ?it/s]  
647 Extracting data files: 100%|██████████| 1/1 [00:00<  
00:00, 135.25it/s]  
648  
    Downloading and preparing dataset csv/default to  
    C:/Users/admin/.cache/huggingface/datasets/csv/  
    default-6c1893227dd4d32a/0.0.0/
```

```
648    6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
       3bcc7ec513e1...  
649 100%|██████████| 1/1 [00:00<00:00, 64.69it/s]  
650 Dataset csv downloaded and prepared to C:/Users/  
       admin/.cache/huggingface/datasets/csv/default-  
       6c1893227dd4d32a/0.0.0/  
       6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
       3bcc7ec513e1. Subsequent calls will reuse this data.  
651 Downloading data files: 100%|██████████| 1/1 [00:00  
       <, ?it/s]  
652 Downloading and preparing dataset csv/default to C:/  
       Users/admin/.cache/huggingface/datasets/csv/default-  
       ff045d55d64fd9ef/0.0.0/  
       6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
       3bcc7ec513e1...  
653 Extracting data files: 100%|██████████| 1/1 [00:00<  
       00:00, 125.47it/s]  
654 100%|██████████| 1/1 [00:00<00:00, 66.96it/s]  
655 Dataset csv downloaded and prepared to C:/Users/  
       admin/.cache/huggingface/datasets/csv/default-  
       ff045d55d64fd9ef/0.0.0/  
       6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
       3bcc7ec513e1. Subsequent calls will reuse this data.  
656 t is : 4  
657 addlen is: 2000  
658 noexp_df.shape[0] last 4605  
659 unexplained dataset length is:  
660 4605  
661 embedding begin  
662 trian and val dataset length: 127530  
663 diversity_sampling_openai  
664 diversity_sampling_openai  
665 diversity_sampling_openai  
666 diversity_sampling_openai  
667 diversity_sampling_openai  
668 diversity_sampling_openai  
669 diversity_sampling_openai  
670 diversity_sampling_openai  
671 981.0  
672 ./embeddings/NEW_bertie_embeddings_textattack/bert-
```

```
672 base-uncased-MNLI_subset_1.pt
673 Found cached dataset csv (C:/Users/admin/.cache/
    huggingface/datasets/csv/default-494a068ef8ae599d/0.
    0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
    3bcc7ec513e1)
674 100%|██████████| 1/1 [00:00<00:00, 52.80it/s]
675 test len 20865
676 Downloading and preparing dataset csv/default to C:/Users/admin/.cache/huggingface/datasets/csv/default-c523cb0ac9895553/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
    3bcc7ec513e1...
677 Downloading data files: 100%|██████████| 1/1 [00:00
    <?, ?it/s]
678 Extracting data files: 100%|██████████| 1/1 [00:00<
    00:00, 167.21it/s]
679 100%|██████████| 1/1 [00:00<00:00, 77.17it/s]
680 Dataset csv downloaded and prepared to C:/Users/
    admin/.cache/huggingface/datasets/csv/default-
    c523cb0ac9895553/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
    3bcc7ec513e1. Subsequent calls will reuse this data.
681 preprocess_samples begin ...
682 uncertainty shape: 1530
683 ./test_embeddings/NEW_bertie_embeddings_textattack/
    bert-base-uncased-MNLI_subset_1.pt
684 classify begin
685 9847
686 0%|          | 0/20 [00:00<?, ?it/s]
    diversity_sampling_openai
687 diversity_sampling_openai
688 1072it [00:16, 489.93it/s]diversity_sampling_openai
689 diversity_sampling_openai
690 diversity_sampling_openai
691 diversity_sampling_openai
692 2165it [00:30, 510.35it/s]diversity_sampling_openai
693 diversity_sampling_openai
694 diversity_sampling_openai
695 diversity_sampling_openai
696 3288it [00:45, 498.03it/s]diversity_sampling_openai
```

```
697 diversity_sampling_openai
698 diversity_sampling_openai
699 3312it [00:55, 498.03it/s]diversity_sampling_openai
700 4414it [00:59, 558.94it/s]diversity_sampling_openai
701 diversity_sampling_openai
702 diversity_sampling_openai
703 diversity_sampling_openai
704 5509it [01:14, 510.33it/s]diversity_sampling_openai
705 diversity_sampling_openai
706 diversity_sampling_openai
707 5520it [01:25, 510.33it/s]diversity_sampling_openai
708 6577it [01:29, 531.07it/s]diversity_sampling_openai
709 diversity_sampling_openai
710 diversity_sampling_openai
711 diversity_sampling_openai
712 7666it [01:43, 512.64it/s]diversity_sampling_openai
713 diversity_sampling_openai
714 diversity_sampling_openai
715 7729it [01:57, 20.80it/s] diversity_sampling_openai
716 8832it [01:58, 585.00it/s]diversity_sampling_openai
717 diversity_sampling_openai
718 diversity_sampling_openai
719 diversity_sampling_openai
720 9935it [02:12, 513.53it/s]diversity_sampling_openai
721 diversity_sampling_openai
722 diversity_sampling_openai
723 diversity_sampling_openai
724 11035it [02:26, 523.79it/s]diversity_sampling_openai
725 diversity_sampling_openai
726 diversity_sampling_openai
727 diversity_sampling_openai
728 12136it [02:41, 509.78it/s]diversity_sampling_openai
729 diversity_sampling_openai
730 diversity_sampling_openai
731 diversity_sampling_openai
732 13224it [02:55, 500.95it/s]diversity_sampling_openai
733 diversity_sampling_openai
734 diversity_sampling_openai
735 diversity_sampling_openai
736 14333it [03:10, 446.56it/s]diversity_sampling_openai
737 diversity_sampling_openai
```

```
738 diversity_sampling_openai
739 diversity_sampling_openai
740 15432it [03:25, 506.75it/s]diversity_sampling_openai
741 diversity_sampling_openai
742 diversity_sampling_openai
743 diversity_sampling_openai
744 16498it [03:40, 504.13it/s]diversity_sampling_openai
745 diversity_sampling_openai
746 diversity_sampling_openai
747 diversity_sampling_openai
748 17648it [03:55, 609.19it/s]diversity_sampling_openai
749 diversity_sampling_openai
750 diversity_sampling_openai
751 diversity_sampling_openai
752 18728it [04:11, 540.72it/s]diversity_sampling_openai
753 diversity_sampling_openai
754 diversity_sampling_openai
755 diversity_sampling_openai
756 19801it [04:26, 495.96it/s]diversity_sampling_openai
757 diversity_sampling_openai
758 diversity_sampling_openai
759 diversity_sampling_openai
760 20938it [04:40, 512.40it/s]diversity_sampling_openai
761 diversity_sampling_openai
762 diversity_sampling_openai
763 diversity_sampling_openai
764 22037it [04:55, 505.56it/s]diversity_sampling_openai
765 diversity_sampling_openai
766 epoch number 19
767 diversity_sampling_openai
768 22080it [05:05, 505.56it/s]diversity_sampling_openai
769 diversity_sampling_openai
770 diversity_sampling_openai
771 global t is : 4
772 test epoch results {'accuracy': 0.5849673202614379
, 'f1_weighted': 0.5589827044381368, 'f1_macro': 0.
48115712586755816, 'f1': array([0.86513995, 0.
26923077, 0.51612903, 0.23602484, 0.58660508,
773 0.09708738, 0.58892128, 0.60861595, 0.
56265985]), 'precision': array([0.85858586, 0.
46666667, 0.54545455, 0.51351351, 0.57990868,
```

```
774     0.35714286, 0.59763314, 0.51956182, 0.  
      56410256]), 'recall': array([0.87179487, 0.18918919  
      , 0.48979592, 0.15322581, 0.59345794,  
775     0.05617978, 0.58045977, 0.73451327, 0.  
      56122449])}  
776 val epoch results {'accuracy': 0.5932721712538226, '  
    f1_weighted': 0.5570437942708063, 'f1_macro': 0.  
    4302162375522831, 'f1': array([0.88963211, 0  
    .          , 0.33333333, 0.29126214, 0.59060403,  
777     0.          , 0.54028436, 0.60289855, 0.  
    62393162]), 'precision': array([0.84177215, 0  
    .          , 0.44444444, 0.71428571, 0.54320988,  
778     0.          , 0.56435644, 0.50980392, 0.  
    66363636]), 'recall': array([0.94326241, 0  
    .          , 0.26666667, 0.18292683, 0.64705882,  
779     0.          , 0.51818182, 0.73758865, 0.  
    58870968])}  
780 22080it [05:11, 70.84it/s]  
781 diversity new embedding 4605  
782 diversity len embedding 4605  
783 MLP_1h(  
784     (l1): Linear(in_features=39, out_features=100,  
      bias=True)  
785     (l2): Linear(in_features=100, out_features=9, bias  
      =True)  
786 )  
787 [2110, 1219, 3094, 2746, 3942, 3263, 4203, 2285,  
  4582, 4580, 4106, 1916, 3159, 3209, 1256, 150, 609,  
  1442, 4461, 2608, 3404, 2010, 372, 2906, 1714, 1990  
  , 1522, 2212, 3342, 3959, 1024, 1130, 3533, 3338,  
  971, 2430, 4088, 1123, 3246, 1661, 407, 1440, 940,  
  2450, 749, 337, 1986, 12, 4519, 1966, 3841, 354,  
  3714, 2361, 381, 259, 4239, 3480, 2316, 1725, 3112,  
  3254, 2962, 2113, 2099, 4414, 4286, 2557, 3069, 217  
  , 1624, 4498, 3909, 3958, 2778, 1958, 2973, 593,  
  2252, 1029, 247, 1565, 1195, 1824, 3548, 887, 2475,  
  4113, 1473, 1723, 3176, 696, 1077, 4125, 432, 4383,  
  448, 677, 3968, 2893, 782, 2309, 2461, 3660, 1692,  
  3072, 2892, 3844, 174, 1043, 3879, 962, 3655, 460,  
  2797, 4430, 3283, 939, 4058, 4540, 4192, 4590, 3129  
  , 1873, 3065, 1143, 263, 1034, 3866, 3322, 231, 1359
```

787 , 1144, 3795, 145, 4261, 297, 412, 678, 2456, 511, 4042, 3751, 3030, 4307, 964, 373, 1926, 990, 3631, 3794, 323, 4494, 3499, 2782, 3550, 925, 4168, 1583, 2389, 3693, 1518, 2696, 2410, 2330, 4260, 1575, 907 , 404, 4277, 294, 375, 116, 2933, 2835, 2613, 384, 1753, 427, 1739, 2129, 2502, 1608, 3153, 2967, 631, 3203, 167, 2048, 1025, 4117, 4147, 1928, 1149, 2235 , 4031, 3624, 2559, 817, 2669, 639, 605, 3658, 301, 3424, 3396, 490, 1589, 1623, 841, 3132, 1417, 1918, 3445, 866, 1031, 2431, 351, 498, 1198, 2124, 2117, 2533, 768, 1471, 758, 2278, 3401, 2939, 3314, 854, 630, 1227, 4360, 3871, 1140, 2222, 1527, 973, 2279, 1405, 3589, 4008, 2336, 4408, 592, 2327, 252, 4523, 1767, 3563, 352, 4189, 4059, 3015, 983, 3358, 1805, 3366, 4022, 3730, 1065, 3823, 3711, 296, 2047, 3788 , 4435, 1305, 3617, 3117, 1441, 3355, 3078, 2841, 3704, 2391, 1960, 4097, 3860, 390, 1167, 2982, 4509 , 2602, 4409, 641, 2534, 72, 1111, 1374, 4116, 2331 , 2440, 2786, 3766, 362, 838, 1108, 317, 4243, 421, 1981, 2818, 2074, 3088, 835, 4579, 1362, 3350, 3466 , 3996, 555, 2648, 1667, 3049, 1992, 3547, 26, 2233 , 2318, 1627, 2302, 369, 919, 634, 818, 2273, 492, 2243, 508, 1206, 824, 945, 3682, 2995, 75, 946, 3392 , 1764, 1858, 1128, 3898, 3465, 3568, 108, 3098, 3505, 2460, 1542, 1817, 1837, 3019, 1049, 4165, 2582 , 1147, 1369, 954, 3517, 1628, 181, 4289, 449, 1911 , 2817, 2103, 172, 3255, 2591, 1940, 2204, 880, 2543 , 890, 4185, 2915, 548, 1897, 2826, 4123, 1428, 3954 , 2106, 3267, 3607, 4242, 2084, 3691, 1301, 701, 2800, 1483, 2154, 2872, 3217, 847, 716, 3748, 2201, 4379, 279, 3453, 2804, 135, 3849, 4393, 3952, 99, 3757, 71, 4040, 4471, 2392, 536, 1668, 2644, 433, 4526, 1015, 3595, 234, 4112, 2284, 78, 3697, 3097, 4357, 137, 2903, 4347, 4182, 2878, 3713, 3383, 4067 , 1980, 3585, 3385, 3637, 101, 3799, 1508, 1803, 236 , 4432, 4149, 568, 2672, 1285, 2736, 4323, 3644, 2646, 774, 3006, 4513, 3147, 4254, 1210, 1822, 1520 , 3273, 233, 3899, 1116, 3443, 611, 3076, 4501, 826 , 2844, 922, 2014, 1533, 2834, 1572, 851, 3586, 366 , 1660, 1384, 4507, 2386, 2603, 3082, 1138, 1788, 114, 393, 2928, 1705, 1699, 2172, 3743, 54, 3324,

787 4062, 1064, 1363, 1556, 2972, 1578, 3250, 58, 1050, 4186, 2720, 865, 2220, 3303, 1120, 226, 1220, 3780, 2144, 1213, 1742, 2680, 2596, 3514, 1599, 3223, 1953 , 1975, 1429, 3946, 3976, 1927, 1203, 2842, 3128, 2000, 2656, 53, 888, 184, 3806, 415, 2821, 216, 1619 , 800, 2577, 1191, 3972, 2041, 3099, 4353, 1883, 2860, 1951, 2358, 3191, 3666, 1160, 4352, 2699, 4327 , 3412, 4389, 1757, 265, 3188, 1665, 3333, 3541, 1184, 2107, 192, 2761, 1463, 2314, 948, 1096, 1260, 720, 1322, 3639, 2469, 2961, 3435, 832, 3349, 1207, 3496, 1380, 1484, 3518, 4170, 411, 2920, 229, 3800, 566, 1984, 2820, 1145, 3694, 401, 4433, 3036, 2850, 278, 1669, 119, 2781, 1467, 220, 1289, 3695, 2111, 1422, 2598, 2466, 4262, 1494, 2943, 4598, 3865, 3845 , 140, 1833, 4095, 1682, 1579, 1688, 191, 2855, 3855 , 2827, 1731, 1612, 8, 882, 651, 1007, 2631, 524, 2709, 3062, 4520, 3265, 3914, 148, 1596, 3136, 1917 , 4451, 3400, 4555, 1465, 986, 211, 2919, 3719, 3467 , 170, 466, 1789, 1998, 4012, 3295, 52, 675, 2025, 63, 1728, 11, 1607, 4565, 3506, 388, 399, 2142, 3576 , 966, 4311, 3771, 2118, 2833, 894, 1354, 477, 2987 , 4380, 3045, 1475, 80, 1361, 2093, 674, 1632, 1941 , 4372, 2287, 5, 1082, 2189, 1807, 1238, 2155, 3950 , 934, 1904, 235, 3791, 2057, 3494, 2981, 2780, 3521 , 4193, 3398, 1117, 4577, 1718, 2094, 455, 2625, 4556, 2415, 1245, 1387, 244, 2197, 4272, 3769, 970, 4124, 3393, 1306, 522, 3012, 2735, 505, 1068, 2114, 844, 4517, 3619, 458, 305, 4151, 312, 3662, 330, 3775, 837, 1560, 1963, 3651, 2083, 2445, 2724, 3643 , 4410, 834, 3648, 2290, 3861, 377, 4044, 1838, 2585 , 697, 419, 3300, 3440, 2012, 1610, 1353, 1246, 2249 , 2768, 652, 2242, 1005, 3457, 3010, 2310, 2108, 3934, 2880, 1449, 658, 1906, 4466, 4481, 4426, 2217 , 2482, 4528, 4220, 3417, 4377, 4118, 1751, 457, 59 , 242, 95, 2856, 3520, 4036, 1454, 4346, 2406, 1600 , 876, 2399, 38, 3796, 1645, 4578, 2620, 1636, 3308 , 770, 710, 1973, 3135, 6, 2632, 1113, 848, 781, 599 , 2351, 822, 4075, 4568, 1564, 2317, 4089, 1887, 1890, 1310, 1458, 3002, 1979, 515, 120, 2485, 3918, 3178, 2301, 3540, 1165, 1584, 4545, 4030, 391, 2542 , 3970, 4016, 3495, 1908, 3059, 440, 130, 1894, 3675

787 , 912, 2251, 306, 2458, 2732, 1261, 1942, 3439, 2282  
, 2181, 1148, 1679, 2209, 185, 2541, 994, 3580, 4446  
, 3852, 4194, 4285, 1106, 3282, 3734, 3138, 3166,  
446, 2612, 4436, 2157, 2635, 3437, 2980, 594, 3107,  
4317, 3838, 1889, 1621, 614, 914, 292, 4552, 1014,  
2953, 2870, 395, 4303, 1434, 3633, 1697, 4032, 131,  
2258, 1396, 4362, 1046, 4274]

788 Label: 4

789 generate a string that only contains the words in  
the text that are only semantically equal to  
donation or offers or volunteering:

790

791 India sends planes with rescuers, essentials to  
Kathmandu http http (photo via Please help spread  
the word. Vanuatu needs our help. No one knows what  
happened. http India offered \$1 Billion loan help  
\$10+ Mil as aid US offer \$1M as aid + Army of  
Missionaries Soul Vultures http £5 can provide a  
family with clean water for a month. Text DEC to  
70000 to donate £5 Nepal Quake http Jammu and  
Kashmir floods: Army has given medical aid to over  
53,000 victims http -via.

792

793

794 AI Explanation: "India offered \$1 Billion loan help  
\$10+ Mil as aid US offer \$1M as aid + donation Soul

795 Downloading and preparing dataset csv/default to C:/  
Users/admin/.cache/huggingface/datasets/csv/default-  
28a81c6decbe8166/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1...

796 Downloading data files: 100%|██████████| 1/1 [00:00  
<, ?it/s]

797 Extracting data files: 100%|██████████| 1/1 [00:00<  
00:00, 166.90it/s]

798 100%|██████████| 1/1 [00:00<00:00, 91.07it/s]

799 Dataset csv downloaded and prepared to C:/Users/  
admin/.cache/huggingface/datasets/csv/default-  
28a81c6decbe8166/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1. Subsequent calls will reuse this data.

```
800 Downloading data files: 100%|██████████| 1/1 [00:00<00:00, 961.56it/s]
801 Extracting data files: 100%|██████████| 1/1 [00:00<00:00, 167.26it/s]
802 Generating train split: 0 examples [00:00, ? examples/s] Downloading and preparing dataset csv/default to C:/Users/admin/.cache/huggingface/datasets/csv/default-390586ad7ea04465/0.0.0/6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1...
803 100%|██████████| 1/1 [00:00<00:00, 52.19it/s]
804 Dataset csv downloaded and prepared to C:/Users/admin/.cache/huggingface/datasets/csv/default-390586ad7ea04465/0.0.0/6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1. Subsequent calls will reuse this data.
805 Downloading data files: 100%|██████████| 1/1 [00:00<?, ?it/s]
806 Downloading and preparing dataset csv/default to C:/Users/admin/.cache/huggingface/datasets/csv/default-1b670dbc819ce91f/0.0.0/6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1...
807 Extracting data files: 100%|██████████| 1/1 [00:00<00:00, 138.61it/s]
808 100%|██████████| 1/1 [00:00<00:00, 62.71it/s]
809 Dataset csv downloaded and prepared to C:/Users/admin/.cache/huggingface/datasets/csv/default-1b670dbc819ce91f/0.0.0/6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1. Subsequent calls will reuse this data.
810 Downloading data files: 100%|██████████| 1/1 [00:00<?, ?it/s]
811 Downloading and preparing dataset csv/default to C:/Users/admin/.cache/huggingface/datasets/csv/default-b017a95e23f242d5/0.0.0/6954658bab30a358235fa864b05cf819af0e179325c740e4bc853bcc7ec513e1...
812 Extracting data files: 100%|██████████| 1/1 [00:00<00:00, 125.41it/s]
813 100%|██████████| 1/1 [00:00<00:00, 111.00it/s]
```

```
814 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
b017a95e23f242d5/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1. Subsequent calls will reuse this data.
815 t is : 5
816 addlen is: 2500
817 noexp_df.shape[0] last 3684
818 unexplained dataset length is:
819 3684
820 embedding begin
821 trian and val dataset length: 149940
822 diversity_sampling_openai
823 diversity_sampling_openai
824 diversity_sampling_openai
825 diversity_sampling_openai
826 diversity_sampling_openai
827 diversity_sampling_openai
828 diversity_sampling_openai
829 diversity_sampling_openai
830 1071.0
831 ./embeddings/NEW_bertie_embeddings_textattack/bert-
base-uncased-MNLI_subset_1.pt
832 Found cached dataset csv (C:/Users/admin/.cache/
huggingface/datasets/csv/default-494a068ef8ae599d/0.
0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1)
833 100%|██████████| 1/1 [00:00<00:00, 64.95it/s]
834 test len 22470
835 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
836 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 143.02it/s]
837 Generating train split: 0 examples [00:00, ?
examples/s]Downloading and preparing dataset csv/
default to C:/Users/admin/.cache/huggingface/
datasets/csv/default-77e40be35ef53c58/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1...
838 100%|██████████| 1/1 [00:00<00:00, 62.71it/s]
```

```
839 Dataset csv downloaded and prepared to C:/Users/
    admin/.cache/huggingface/datasets/csv/default-
    77e40be35ef53c58/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
    3bcc7ec513e1. Subsequent calls will reuse this data.
840 preprocess_samples begin ...
841 uncertainty shape: 1530
842 ./test_embeddings/NEW_bertie_embeddings_textattack/
    bert-base-uncased-MNLI_subset_1.pt
843 classify begin
844 10768
845 0%|          | 0/20 [00:00<?, ?it/s]
    diversity_sampling_openai
846 diversity_sampling_openai
847 1170it [00:09, 728.80it/s]diversity_sampling_openai
848 diversity_sampling_openai
849 diversity_sampling_openai
850 diversity_sampling_openai
851 2359it [00:23, 572.27it/s]diversity_sampling_openai
852 diversity_sampling_openai
853 diversity_sampling_openai
854 diversity_sampling_openai
855 3556it [00:38, 544.82it/s]diversity_sampling_openai
856 diversity_sampling_openai
857 diversity_sampling_openai
858 diversity_sampling_openai
859 4757it [00:52, 569.52it/s]diversity_sampling_openai
860 diversity_sampling_openai
861 diversity_sampling_openai
862 4820it [01:03, 569.52it/s]diversity_sampling_openai
863 5982it [01:08, 551.98it/s]diversity_sampling_openai
864 diversity_sampling_openai
865 diversity_sampling_openai
866 diversity_sampling_openai
867 7202it [01:22, 578.26it/s]diversity_sampling_openai
868 diversity_sampling_openai
869 diversity_sampling_openai
870 7230it [01:33, 578.26it/s]diversity_sampling_openai
871 8385it [01:38, 457.90it/s]diversity_sampling_openai
872 diversity_sampling_openai
873 diversity_sampling_openai
```

```
874 diversity_sampling_openai
875 9632it [01:52, 640.55it/s]diversity_sampling_openai
876 diversity_sampling_openai
877 diversity_sampling_openai
878 9641it [02:05, 16.63it/s] diversity_sampling_openai
879 10790it [02:07, 616.05it/s]diversity_sampling_openai
880 diversity_sampling_openai
881 diversity_sampling_openai
882 diversity_sampling_openai
883 11961it [02:21, 576.70it/s]diversity_sampling_openai
884 diversity_sampling_openai
885 diversity_sampling_openai
886 12050it [02:33, 576.70it/s]diversity_sampling_openai
887 13239it [02:36, 584.26it/s]diversity_sampling_openai
888 diversity_sampling_openai
889 diversity_sampling_openai
890 diversity_sampling_openai
891 14407it [02:51, 542.51it/s]diversity_sampling_openai
892 diversity_sampling_openai
893 diversity_sampling_openai
894 14460it [03:03, 542.51it/s]diversity_sampling_openai
895 15622it [03:06, 561.67it/s]diversity_sampling_openai
896 diversity_sampling_openai
897 diversity_sampling_openai
898 diversity_sampling_openai
899 16846it [03:21, 565.37it/s]diversity_sampling_openai
900 diversity_sampling_openai
901 diversity_sampling_openai
902 16871it [03:34, 13.94it/s] diversity_sampling_openai
903 18047it [03:36, 573.64it/s]diversity_sampling_openai
904 diversity_sampling_openai
905 diversity_sampling_openai
906 diversity_sampling_openai
907 19193it [03:51, 574.55it/s]diversity_sampling_openai
908 diversity_sampling_openai
909 diversity_sampling_openai
910 19281it [04:04, 22.34it/s] diversity_sampling_openai
911 20443it [04:05, 627.28it/s]diversity_sampling_openai
912 diversity_sampling_openai
913 diversity_sampling_openai
914 diversity_sampling_openai
```

```
915 21601it [04:19, 568.86it/s]diversity_sampling_openai
916 diversity_sampling_openai
917 diversity_sampling_openai
918 diversity_sampling_openai
919 22845it [04:34, 557.41it/s]diversity_sampling_openai
920 diversity_sampling_openai
921 diversity_sampling_openai
922 diversity_sampling_openai
923 24061it [04:48, 588.81it/s]diversity_sampling_openai
924 diversity_sampling_openai
925 epoch number 19
926 diversity_sampling_openai
927 diversity_sampling_openai
928 24100it [05:03, 588.81it/s]diversity_sampling_openai
929 diversity_sampling_openai
930 global t is : 5
931 test epoch results {'accuracy': 0.5908496732026144
, 'f1_weighted': 0.5713345578711213, 'f1_macro': 0.
5088986688834021, 'f1': array([0.86848635, 0.
32786885, 0.5625 , 0.35087719, 0.56202532,
932 0.14159292, 0.61538462, 0.60924751, 0.
54210526]), 'precision': array([0.84134615, 0.
41666667, 0.57446809, 0.63829787, 0.61325967,
933 0.33333333, 0.63414634, 0.51612903, 0.
55978261]), 'recall': array([0.8974359 , 0.27027027
, 0.55102041, 0.24193548, 0.51869159,
934 0.08988764, 0.59770115, 0.74336283, 0.5255102
])}
935 val epoch results {'accuracy': 0.5863678804855276, 'f1_weighted': 0.5629589655206216, 'f1_macro': 0.
4641962282476476, 'f1': array([0.86322188, 0.
11764706, 0.37735849, 0.40298507, 0.53818182,
936 0.17391304, 0.54393305, 0.61721854, 0.
54330709]), 'precision': array([0.80225989, 0.2
, 0.5 , 0.61363636, 0.57364341,
937 0.54545455, 0.55084746, 0.52242152, 0.
59482759]), 'recall': array([0.93421053, 0.08333333
, 0.3030303 , 0.3 , 0.50684932,
938 0.10344828, 0.53719008, 0.75404531, 0.5
])}
939 24100it [05:05, 78.93it/s]
```

```
940 diversity new embedding 3684
941 diversity len embedding 3684
942 MLP_1h(
943     (l1): Linear(in_features=42, out_features=100,
944         bias=True)
944     (l2): Linear(in_features=100, out_features=9, bias
945         =True)
945 )
946 [3434, 3582, 2830, 267, 2599, 870, 481, 2949, 2411,
947 422, 95, 2243, 1509, 2989, 1540, 2113, 3265, 275,
948 3572, 3295, 854, 860, 2008, 30, 3098, 1300, 2421,
949 2293, 1497, 2982, 149, 3365, 347, 2980, 1503, 1992,
950 2315, 2035, 178, 87, 705, 3493, 3424, 631, 3152,
951 2291, 2767, 242, 1015, 1089, 782, 2998, 46, 3559,
952 897, 284, 1648, 3289, 2740, 2921, 55, 2126, 2810,
953 2331, 3026, 562, 3348, 1805, 1998, 2602, 572, 2624,
954 318, 2967, 3528, 3518, 1379, 967, 2812, 381, 455, 98
955 , 1647, 2121, 487, 2855, 2852, 1002, 1265, 2455, 817
956 , 326, 2924, 2603, 1796, 1752, 781, 31, 2817, 3629,
957 3009, 424, 1776, 3282, 2083, 3380, 3413, 3059, 1230
958 , 1561, 1964, 2592, 3345, 2022, 1950, 1205, 2728,
959 821, 2521, 2246, 2820, 307, 2149, 1196, 1432, 3121,
960 3251, 2860, 729, 6, 2145, 1761, 2841, 1466, 2392,
961 1314, 1438, 2052, 1833, 709, 3517, 2381, 2261, 2460
962 , 2245, 976, 1312, 1421, 1022, 2691, 1719, 917, 48,
963 2970, 1575, 510, 2649, 1112, 82, 2747, 3047, 619,
964 935, 2338, 1504, 2566, 1548, 2862, 2434, 3543, 1718
965 , 2861, 1353, 3015, 3516, 2546, 164, 1921, 1189,
966 3225, 1868, 299, 798, 262, 1105, 893, 621, 1742,
967 1299, 3333, 1615, 846, 492, 2666, 2979, 2591, 2010,
968 2900, 2783, 3568, 1745, 1652, 1883, 973, 3192, 1624
969 , 1743, 1307, 910, 100, 699, 2087, 594, 1243, 2772,
970 1513, 2876, 2621, 1420, 3115, 405, 3177, 2471, 722,
971 3066, 2407, 3321, 2987, 2326, 3287, 2473, 42, 1318,
972 1298, 2319, 350, 3072, 2709, 3595, 1129, 248, 3257,
973 3482, 2722, 1870, 2301, 1851, 3303, 3344, 2577, 28,
974 936, 1272, 2345, 2946, 3111, 1881, 3035, 2983, 2353
975 , 3429, 1640, 2307, 96, 2158, 3385, 1538, 1311, 3581
976 , 1185, 979, 1116, 1293, 2241, 2585, 2945, 2018,
977 2047, 2360, 3383, 462, 2450, 2475, 170, 3597, 3430,
978 2014, 1207, 475, 5, 172, 1404, 1612, 3404, 1210,
```

|     |   |
|-----|---|
| 946 | 2623, 2009, 1906, 165, 2565, 3510, 1133, 1871, 3504<br>, 3062, 2023, 3226, 2729, 3180, 2248, 2235, 623, 733<br>, 213, 1616, 2699, 2234, 3162, 89, 964, 1603, 2775,<br>2576, 2554, 644, 255, 3651, 1254, 3387, 2025, 3487,<br>3323, 3623, 3624, 54, 2906, 1611, 614, 538, 660,<br>2034, 1029, 1770, 36, 1310, 1798, 3216, 3054, 701,<br>695, 1812, 494, 3679, 2321, 2850, 3023, 2425, 1668,<br>123, 2356, 2495, 401, 1636, 2914, 2403, 1326, 3635,<br>1608, 2, 996, 585, 1167, 2485, 2616, 681, 358, 2502<br>, 1441, 2221, 1885, 3354, 1412, 1366, 3462, 2296,<br>837, 3170, 617, 1018, 3448, 3465, 2553, 2925, 2662,<br>63, 653, 749, 843, 845, 1199, 1613, 2960, 1715, 402<br>, 2529, 3053, 72, 2279, 1807, 3031, 114, 3077, 2943<br>, 2931, 1928, 17, 1905, 1235, 2347, 115, 3674, 2424<br>, 1705, 2375, 370, 1456, 1605, 2352, 1983, 1046,<br>2522, 2604, 2285, 1610, 3410, 2763, 3649, 442, 1270<br>, 3655, 1074, 437, 1977, 1281, 1597, 2573, 2984, 393<br>, 1590, 3132, 124, 3563, 3439, 1791, 1800, 2667, 873<br>, 1975, 2514, 566, 676, 2290, 183, 1329, 861, 118,<br>3631, 1840, 2893, 1913, 577, 992, 1128, 1321, 1825,<br>634, 2143, 3267, 1827, 2771, 3148, 2415, 1169, 2695<br>, 229, 2437, 3592, 1551, 885, 3275, 2660, 2037, 1417<br>, 1040, 1240, 1834, 217, 1642, 1317, 2050, 2259,<br>3452, 1157, 3, 3294, 21, 2644, 225, 2297, 2013, 3342<br>, 1060, 1949, 2287, 3587, 2446, 2868, 2300, 2560,<br>1584, 1672, 3223, 1582, 3205, 800, 3553, 1951, 245,<br>1907, 1915, 3670, 3409, 750, 543, 839, 3140, 274,<br>340, 2954, 2958, 3100, 1309, 1867, 3224, 2105, 592,<br>3231, 2749, 2909, 3566, 2210, 467, 3463, 3052, 808,<br>3002, 907, 1952, 1746, 3669, 2085, 2559, 2923, 2114<br>, 1352, 59, 1810, 789, 590, 1760, 1166, 2271, 33,<br>2498, 649, 1319, 1496, 822, 1013, 753, 2995, 3074,<br>2383, 3678, 2580, 1944, 1491, 2172, 2880, 1917, 260<br>, 1058, 1516, 2858, 3598, 905, 3343, 1547, 496, 934<br>, 1579, 1442, 3548, 1723, 1135, 1831, 1830, 1924,<br>876, 3127, 1389, 1255, 107, 3034, 2177, 86, 497,<br>2587, 2733, 2099, 1091, 3457, 23, 2000, 2431, 2372,<br>3509, 686, 1987, 1308, 2750, 1599, 1923, 293, 194,<br>1670, 2081, 1095, 3584, 2274, 1053, 3529, 1893, 2952<br>, 1607, 3298, 948, 3073, 324, 2174, 3508, 1967, 583<br>, 3652, 412, 1645, 2169, 3281, 3211, 75, 202, 1477, |
|-----|---|

946 1933, 18, 1140, 3552, 420, 1458, 3325, 1664, 1900,  
1212, 683, 3240, 8, 1098, 847, 3600, 3416, 3645,  
2519, 3202, 735, 922, 49, 3008, 2780, 3266, 1041,  
466, 12, 2787, 1170, 2520, 2432, 2881, 2901, 360,  
1073, 1526, 809, 1720, 2992, 1573, 3422, 2799, 2458  
, 1854, 3367, 3486, 1238, 2898, 312, 480, 3667, 2024  
, 1872, 2395, 3625, 3606, 1384, 1644, 2314, 1661,  
3460, 418, 90, 666, 1422, 2827, 1494, 1055, 2633,  
3080, 2292, 2622]

947 Label: 5

948 generate a string that only contains the words in  
the text that are only semantically equal to caution  
and advice:

949

950 Tsunami warning is now in effect for Chile and Peru  
, following the 7.8 quake on Wednesday http nussa! "  
Climate change is intensifying the risks for  
hundreds of millions of people" - Ban http Pam http  
"@htTweets: Chile: Massive 8.2 earthquake hits the  
coast; sparks tsunami http Hagupit is now a Super  
Typhoon and looks to reach maximum potential  
intensity. Philippines: prepare for landfall &;  
focu... Thanks partly to developing El Niño, Typhoon  
Hagupit will be weaker than Hayan at Philippine  
landfall. http.

951

952

953 AI Explanation: "Caution advisory Tsunami warning  
now effect Chile Peru following quake Wednesday  
Climate change intensifying risks hundreds millions  
people prepare

954 Downloading and preparing dataset csv/default to C:/  
Users/admin/.cache/huggingface/datasets/csv/default-  
def81b5282d3b449/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85  
3bcc7ec513e1...

955 Downloading data files: 100%|██████████| 1/1 [00:00  
<, ?it/s]

956 Extracting data files: 100%|██████████| 1/1 [00:00<  
00:00, 143.31it/s]

957 100%|██████████| 1/1 [00:00<00:00, 76.84it/s]

```
958 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
def81b5282d3b449/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1. Subsequent calls will reuse this data.
959 Downloading data files: 100%|██████████| 1/1 [00:00<
00:00, 950.01it/s]
960 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 168.79it/s]
961 Generating train split: 0 examples [00:00, ?
examples/s]Downloading and preparing dataset csv/
default to C:/Users/admin/.cache/huggingface/
datasets/csv/default-1650d5b0f75e40f8/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1...
962 100%|██████████| 1/1 [00:00<00:00, 62.71it/s]
963 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
1650d5b0f75e40f8/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1. Subsequent calls will reuse this data.
964 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
965 Downloading and preparing dataset csv/default to C:/
Users/admin/.cache/huggingface/datasets/csv/default-
38a21a148da4e86b/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1...
966 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 134.20it/s]
967 100%|██████████| 1/1 [00:00<00:00, 77.23it/s]
968 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
38a21a148da4e86b/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc85
3bcc7ec513e1. Subsequent calls will reuse this data.
969 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
970 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 142.24it/s]
971 Downloading and preparing dataset csv/default to C:/
```

```
971 Users/admin/.cache/huggingface/datasets/csv/default  
-2f9a3e94ef8d5f6b/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8  
53bcc7ec513e1...  
972 100%|██████████| 1/1 [00:00<00:00, 76.75it/s]  
973 Dataset csv downloaded and prepared to C:/Users/  
admin/.cache/huggingface/datasets/csv/default-  
2f9a3e94ef8d5f6b/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8  
53bcc7ec513e1. Subsequent calls will reuse this  
data.  
974 t is : 6  
975 addlen is: 3000  
976 noexp_df.shape[0] last 2948  
977 unexplained dataset length is:  
978 2948  
979 embedding begin  
980 trian and val dataset length: 171450  
981 diversity_sampling_openai  
982 diversity_sampling_openai  
983 diversity_sampling_openai  
984 diversity_sampling_openai  
985 diversity_sampling_openai  
986 diversity_sampling_openai  
987 diversity_sampling_openai  
988 diversity_sampling_openai  
989 1143.0  
990 ./embeddings/NEW_bertie_embeddings_textattack/bert-  
base-uncased-MNLI_subset_1.pt  
991 Found cached dataset csv (C:/Users/admin/.cache/  
huggingface/datasets/csv/default-494a068ef8ae599d/0  
.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8  
53bcc7ec513e1)  
992 100%|██████████| 1/1 [00:00<00:00, 402.64it/s]  
993 test len 24075  
994 Downloading data files: 100%|██████████| 1/1 [00:00  
<, ?it/s]  
995 Extracting data files: 100%|██████████| 1/1 [00:00<  
00:00, 110.77it/s]  
996 Downloading and preparing dataset csv/default to C
```

```
996 :/Users/admin/.cache/huggingface/datasets/csv/
      default-75faad87142d3ba3/0.0.0/
      6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
      53bcc7ec513e1...
997 100%|██████████| 1/1 [00:00<00:00, 83.47it/s]
998 Dataset csv downloaded and prepared to C:/Users/
      admin/.cache/huggingface/datasets/csv/default-
      75faad87142d3ba3/0.0.0/
      6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
      53bcc7ec513e1. Subsequent calls will reuse this
      data.
999 preprocess_samples begin ...
1000 uncertainty shape: 1530
1001 ./test_embeddings/NEW_bertie_embeddings_textattack/
      bert-base-uncased-MNLI_subset_1.pt
1002 classify begin
1003 11504
1004    0%|           | 0/20 [00:00<?, ?it/s]
      diversity_sampling_openai
1005 diversity_sampling_openai
1006 1270it [00:12, 582.90it/s]diversity_sampling_openai
1007 diversity_sampling_openai
1008 1286it [00:22, 582.90it/s]diversity_sampling_openai
1009 diversity_sampling_openai
1010 2567it [00:29, 639.28it/s]diversity_sampling_openai
1011 diversity_sampling_openai
1012 diversity_sampling_openai
1013 2573it [00:43, 10.98it/s] diversity_sampling_openai
1014 3827it [00:45, 567.06it/s]diversity_sampling_openai
1015 diversity_sampling_openai
1016 diversity_sampling_openai
1017 diversity_sampling_openai
1018 5098it [01:01, 454.64it/s]diversity_sampling_openai
1019 diversity_sampling_openai
1020 diversity_sampling_openai
1021 5145it [01:15, 13.27it/s] diversity_sampling_openai
1022 6387it [01:17, 669.07it/s]diversity_sampling_openai
1023 diversity_sampling_openai
1024 diversity_sampling_openai
1025 diversity_sampling_openai
1026 7677it [01:32, 547.66it/s]diversity_sampling_openai
```

```
1027 diversity_sampling_openai
1028 diversity_sampling_openai
1029 7717it [01:46, 13.46it/s] diversity_sampling_openai
1030 8984it [01:48, 576.85it/s] diversity_sampling_openai
1031 diversity_sampling_openai
1032 diversity_sampling_openai
1033 9003it [02:03, 10.78it/s] diversity_sampling_openai
1034 10270it [02:05, 508.40it/s]
    diversity_sampling_openai
1035 diversity_sampling_openai
1036 diversity_sampling_openai
1037 diversity_sampling_openai
1038 11556it [02:22, 527.06it/s]
    diversity_sampling_openai
1039 diversity_sampling_openai
1040 diversity_sampling_openai
1041 11574it [02:32, 527.06it/s]
    diversity_sampling_openai
1042 12847it [02:38, 569.11it/s]
    diversity_sampling_openai
1043 diversity_sampling_openai
1044 diversity_sampling_openai
1045 diversity_sampling_openai
1046 14146it [02:54, 518.13it/s]
    diversity_sampling_openai
1047 diversity_sampling_openai
1048 diversity_sampling_openai
1049 diversity_sampling_openai
1050 15429it [03:09, 654.96it/s]
    diversity_sampling_openai
1051 diversity_sampling_openai
1052 diversity_sampling_openai
1053 15432it [03:22, 654.96it/s]
    diversity_sampling_openai
1054 16678it [03:25, 558.20it/s]
    diversity_sampling_openai
1055 diversity_sampling_openai
1056 diversity_sampling_openai
1057 diversity_sampling_openai
1058 18000it [03:40, 716.01it/s]
    diversity_sampling_openai
```

```
1059 diversity_sampling_openai
1060 diversity_sampling_openai
1061 18005it [03:53, 16.12it/s]
    diversity_sampling_openai
1062 19267it [03:55, 678.65it/s]
    diversity_sampling_openai
1063 diversity_sampling_openai
1064 diversity_sampling_openai
1065 diversity_sampling_openai
1066 20482it [04:10, 606.47it/s]
    diversity_sampling_openai
1067 diversity_sampling_openai
1068 diversity_sampling_openai
1069 20577it [04:24, 21.79it/s]
    diversity_sampling_openai
1070 21807it [04:26, 534.01it/s]
    diversity_sampling_openai
1071 diversity_sampling_openai
1072 diversity_sampling_openai
1073 diversity_sampling_openai
1074 23095it [04:41, 531.92it/s]
    diversity_sampling_openai
1075 diversity_sampling_openai
1076 diversity_sampling_openai
1077 23149it [04:54, 14.71it/s]
    diversity_sampling_openai
1078 24398it [04:56, 579.85it/s]
    diversity_sampling_openai
1079 diversity_sampling_openai
1080 diversity_sampling_openai
1081 diversity_sampling_openai
1082 25677it [05:12, 570.57it/s]
    diversity_sampling_openai
1083 diversity_sampling_openai
1084 epoch number 19
1085 diversity_sampling_openai
1086 25720it [05:22, 570.57it/s]
    diversity_sampling_openai
1087 diversity_sampling_openai
1088 diversity_sampling_openai
1089 global t is : 6
```

```

1090 test epoch results {'accuracy': 0.6071895424836601
, 'f1_weighted': 0.5882673410059415, 'f1_macro': 0.
5292195807865946, 'f1': array([0.86528497, 0.4
, 0.57446809, 0.36781609, 0.59728507,
1091 0.15238095, 0.59710145, 0.62730627, 0.
58133333]), 'precision': array([0.87434555, 0.
61111111, 0.6 , 0.64 , 0.57894737,
1092 0.5 , 0.60233918, 0.53797468, 0.
60893855]), 'recall': array([0.85641026, 0.2972973
, 0.55102041, 0.25806452, 0.61682243,
1093 0.08988764, 0.59195402, 0.75221239, 0.
55612245])}
1094 val epoch results {'accuracy': 0.5888013998250219
, 'f1_weighted': 0.5660316027034559, 'f1_macro': 0.
4708009580591522, 'f1': array([0.90184049, 0.
24390244, 0.37037037, 0.35036496, 0.48484848,
1095 0.08450704, 0.56573705, 0.62025316, 0.
61538462]), 'precision': array([0.88023952, 0.3125
, 0.52631579, 0.57142857, 0.46511628,
1096 0.375 , 0.58677686, 0.5326087 , 0.
63768116]), 'recall': array([0.9245283 , 0.2
, 0.28571429, 0.25263158, 0.50632911,
1097 0.04761905, 0.54615385, 0.74242424, 0.
59459459])}
1098 diversity new embedding 2948
1099 diversity len embedding 2948
1100 MLP_1h(
1101 (l1): Linear(in_features=45, out_features=100,
bias=True)
1102 (l2): Linear(in_features=100, out_features=9,
bias=True)
1103 )
1104 25720it [05:29, 78.11it/s]
1105 [2232, 1886, 665, 89, 2110, 1076, 965, 1609, 1568,
2196, 2633, 1077, 2947, 1342, 2359, 2209, 1291,
2254, 1343, 2582, 1633, 536, 2343, 1813, 762, 909,
1328, 1253, 1323, 2350, 2791, 2379, 418, 1479, 1125
, 2893, 1587, 1462, 945, 1963, 889, 2813, 315, 2774
, 921, 88, 2373, 674, 352, 1833, 2433, 924, 514,
1025, 2279, 2414, 583, 1391, 124, 2581, 809, 1172,
838, 544, 1837, 1004, 374, 2472, 369, 2141, 2365,
```

|      |  |
|------|--|
| 1105 | 751, 1020, 1846, 2824, 1371, 19, 1627, 745, 1297, 176, 2445, 1039, 1022, 1239, 2465, 1758, 1599, 1250, 324, 1406, 41, 1248, 1899, 1394, 28, 1148, 2151, 2659, 1503, 2009, 1829, 2283, 1551, 213, 2114, 1249, 2019, 294, 2658, 2475, 763, 595, 2734, 1261, 1704, 251, 1579, 1967, 395, 367, 2513, 1777, 651, 2717, 1590, 405, 2081, 1049, 1455, 2029, 992, 2720, 2146, 1818, 2602, 2117, 2559, 2577, 1663, 1793, 673, 318, 2567, 2345, 2765, 1037, 1213, 2665, 159, 667, 2617, 1998, 2620, 1184, 362, 1635, 2530, 2394, 785, 2444, 1308, 304, 1741, 975, 429, 972, 1669, 110, 1237, 1090, 686, 885, 1435, 1778, 2594, 1953, 2164, 608, 2124, 1996, 2795, 2263, 1197, 1461, 1350, 2486, 2383, 2677, 2390, 1246, 1766, 293, 937, 925, 68, 2640, 2541, 2490, 1811, 633, 40, 2014, 452, 1227, 827, 948, 1742, 195, 1709, 2729, 1316, 1335, 2314, 2287, 1729, 72, 1087, 2532, 1581, 2618, 2372, 688, 1745, 1486, 2284, 417, 403, 1836, 2001, 1576, 1338, 2271, 317, 344, 851, 1258, 1463, 2572, 766, 2850, 735, 2409, 1108, 1547, 412, 500, 1095, 646, 846, 2446, 2091, 1786, 2460, 1616, 1428, 752, 346, 2826, 279, 2642, 808, 65, 38, 2607, 2905, 1538, 1199, 2181, 1047, 2037, 396, 750, 2025, 853, 879, 1981, 2251, 1143, 2138, 1119, 2549, 519, 2332, 2673, 1557, 717, 2082, 2788, 10, 512, 1894, 1112, 877, 1441, 1637, 2827, 504, 311, 1589, 1731, 1084, 1097, 33, 2591, 1776, 811, 2728, 2191, 792, 284, 1268, 451, 1194, 150, 1540, 537, 1033, 2766, 1629, 1687, 1890, 2926, 630, 2301, 1376, 171, 1493, 552, 1554, 2277, 694, 2657, 1923, 384, 2245, 783, 2423, 2783, 2224, 597, 2366, 2866, 1390, 2479, 1234, 2192, 589, 1266, 108, 2829, 423, 1353, 721, 2352, 1, 42, 2195, 2888, 234, 541, 379, 457, 1499, 1405, 2583, 1178, 1516, 2040, 6, 2632, 1186, 1044, 283, 1496, 138, 895, 2923, 2637, 2052, 297, 225, 243, 1193, 571, 1034, 1652, 1719, 574, 2557, 2855, 2311, 1761, 2927, 1014, 2592, 168, 2122, 556, 1102, 2547, 1987, 951, 107, 2071, 1922, 387, 2194, 1163, 2427, 2517, 1158, 123, 308, 822, 2900, 2108, 663, 2073, 1510, 2305, 1438, 354, 1764, 175, 2522, 90, 2370, 406, 850, 182, 1375, 44, 2727, 94, 727, 1035, 2520 |
|------|--|

1105 , 55, 1521, 1802, 677, 2231, 1876, 1129, 2149, 2079  
, 2624, 1692, 1785, 2403, 2548, 1826, 1607, 100,  
1466, 2576, 1357, 1429, 697, 2539, 2733, 716, 923,  
2843, 1980, 621, 638, 749, 2481, 616, 1161, 212,  
2694, 2831, 1725, 158, 844, 2010, 27, 942, 1041,  
2007, 800, 2878, 1936, 2411, 1593, 73, 620, 1928,  
1272, 1437, 141, 2086, 1151, 1871, 1043, 679, 2315  
, 2275, 194, 1559, 1555, 1275, 2218, 1094, 2185,  
1752, 1807, 1970, 2705, 564, 1740, 528, 2421, 557,  
1561, 1691, 1695, 2684, 106, 2804, 335, 1812, 2864  
, 1523, 759, 629, 1684, 2846, 314, 2767, 296, 353,  
2145, 445, 9, 1063, 1916, 1399, 2362, 1759, 1586,  
954, 2131, 1517, 1018, 1718, 2815, 880, 2152, 102,  
1585, 1658, 1276, 22, 15, 2064, 1880, 459, 2489,  
2062, 1893, 2041, 2570, 676, 2519, 286, 600, 1418,  
2575, 604, 2543, 97, 383, 1969, 1564, 2780, 1783,  
705, 540, 882, 2579, 693, 884, 2374, 1003, 1290]  
1106 Label: 6  
1107 generate a string that only contains the words in  
the text that are only semantically equal to  
sympathy and emotional support:  
1108  
1109 Thinking of people in the Philippines as Typhoon  
Hagupit approaches. As it nears has staff + medical  
a... http P R A Y F O R C H I L E AND PRAY FOR THE  
CENA FAMILY R I P J O H N C E N A http The Diocese  
of Lancaster : Pope prays for victims of massive  
earthquake in Nepal http Guys northern chile really  
needs a support message Pray For Chile Between  
Baltimore and Nepal...I'm emotionally overwhelmed..  
1110  
1111  
1112 AI Explanation: "Sympathy and emotional support  
thinking people Philippines Typhoon Hagupit staff  
medical PRAY FOR CHILE PRAY FOR  
1113 Downloading data files: 100%|██████████| 1/1 [00:00  
<, ?it/s]  
1114 Extracting data files: 100%|██████████| 1/1 [00:00<  
00:00, 157.25it/s]  
1115 Generating train split: 0 examples [00:00, ?  
examples/s]Downloading and preparing dataset csv/

```
1115 default to C:/Users/admin/.cache/huggingface/
datasets/csv/default-6d2f85edd0cd5f06/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1...
1116 100%|██████████| 1/1 [00:00<00:00, 80.54it/s]
1117 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
6d2f85edd0cd5f06/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1. Subsequent calls will reuse this
data.
1118 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
1119 Downloading and preparing dataset csv/default to C
:/Users/admin/.cache/huggingface/datasets/csv/
default-1c299d5326c1a509/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1...
1120 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 167.61it/s]
1121 100%|██████████| 1/1 [00:00<00:00, 66.67it/s]
1122 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
1c299d5326c1a509/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1. Subsequent calls will reuse this
data.
1123 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
1124 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 166.03it/s]
1125 Generating train split: 0 examples [00:00, ?
examples/s]Downloading and preparing dataset csv/
default to C:/Users/admin/.cache/huggingface/
datasets/csv/default-feec121c28961809/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1...
1126 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
feec121c28961809/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
```

```
1126 53bcc7ec513e1. Subsequent calls will reuse this
      data.
1127 100%|██████████| 1/1 [00:00<00:00, 117.31it/s]
1128 Downloading data files: 100%|██████████| 1/1 [00:00
      <, ?it/s]
1129 Extracting data files: 100%|██████████| 1/1 [00:00<
      00:00, 158.87it/s]
1130 Downloading and preparing dataset csv/default to C
      :/Users/admin/.cache/huggingface/datasets/csv/
      default-240a3d02d493343e/0.0.0/
      6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
      53bcc7ec513e1...
1131 100%|██████████| 1/1 [00:00<00:00, 110.96it/s]
1132 Dataset csv downloaded and prepared to C:/Users/
      admin/.cache/huggingface/datasets/csv/default-
      240a3d02d493343e/0.0.0/
      6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
      53bcc7ec513e1. Subsequent calls will reuse this
      data.
1133 t is : 7
1134 addlen is: 3500
1135 noexp_df.shape[0] last 2359
1136 unexplained dataset length is:
1137 2359
1138 embedding begin
1139 trian and val dataset length: 192960
1140 diversity_sampling_openai
1141 diversity_sampling_openai
1142 diversity_sampling_openai
1143 diversity_sampling_openai
1144 diversity_sampling_openai
1145 diversity_sampling_openai
1146 diversity_sampling_openai
1147 diversity_sampling_openai
1148 1206.0
1149 ./embeddings/NEW_bertie_embeddings_textattack/bert-
      base-uncased-MNLI_subset_1.pt
1150 Found cached dataset csv (C:/Users/admin/.cache/
      huggingface/datasets/csv/default-494a068ef8ae599d/0
      .0.0/
      6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
```

```
1150 53bcc7ec513e1)
1151 100%|██████████| 1/1 [00:00<00:00, 59.02it/s]
1152 test len 25680
1153 Downloading and preparing dataset csv/default to C
:/Users/admin/.cache/huggingface/datasets/csv/
default-f98fb13b4bad73b8/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1...
1154 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
1155 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 123.09it/s]
1156 100%|██████████| 1/1 [00:00<00:00, 77.19it/s]
1157 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
f98fb13b4bad73b8/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1. Subsequent calls will reuse this
data.
1158 preprocess_samples begin ...
1159 uncertainty shape: 1530
1160 ./test_embeddings/NEW_bertie_embeddings_textattack/
bert-base-uncased-MNLI_subset_1.pt
1161 classify begin
1162 12093
1163 0%|          | 0/20 [00:00<?, ?it/s]
    diversity_sampling_openai
1164 diversity_sampling_openai
1165 1344it [00:22, 332.99it/s]diversity_sampling_openai
1166 diversity_sampling_openai
1167 diversity_sampling_openai
1168 1358it [00:38, 6.15it/s] diversity_sampling_openai
1169 2658it [00:42, 474.57it/s]diversity_sampling_openai
1170 diversity_sampling_openai
1171 diversity_sampling_openai
1172 2715it [00:57, 10.89it/s] diversity_sampling_openai
1173 4065it [01:02, 212.46it/s]diversity_sampling_openai
1174 diversity_sampling_openai
1175 diversity_sampling_openai
1176 4072it [01:19, 3.86it/s] diversity_sampling_openai
1177 5418it [01:24, 345.79it/s]diversity_sampling_openai
```

```
1178 diversity_sampling_openai
1179 diversity_sampling_openai
1180 5429it [01:38, 7.36it/s] diversity_sampling_openai
1181 6758it [01:42, 424.92it/s] diversity_sampling_openai
1182 diversity_sampling_openai
1183 diversity_sampling_openai
1184 diversity_sampling_openai
1185 8059it [01:57, 701.74it/s] diversity_sampling_openai
1186 diversity_sampling_openai
1187 diversity_sampling_openai
1188 diversity_sampling_openai
1189 9465it [02:13, 396.03it/s] diversity_sampling_openai
1190 diversity_sampling_openai
1191 diversity_sampling_openai
1192 9500it [02:26, 15.46it/s] diversity_sampling_openai
1193 10807it [02:28, 673.00it/s]
    diversity_sampling_openai
1194 diversity_sampling_openai
1195 diversity_sampling_openai
1196 10857it [02:52, 9.92it/s]
    diversity_sampling_openai
1197 12212it [02:55, 243.62it/s]
    diversity_sampling_openai
1198 diversity_sampling_openai
1199 12213it [03:05, 243.62it/s]
    diversity_sampling_openai
1200 diversity_sampling_openai
1201 13542it [03:15, 291.53it/s]
    diversity_sampling_openai
1202 diversity_sampling_openai
1203 13570it [03:25, 291.53it/s]
    diversity_sampling_openai
1204 diversity_sampling_openai
1205 14914it [03:36, 429.55it/s]
    diversity_sampling_openai
1206 diversity_sampling_openai
1207 diversity_sampling_openai
1208 diversity_sampling_openai
1209 16239it [03:53, 519.78it/s]
    diversity_sampling_openai
1210 diversity_sampling_openai
```

```
1211 diversity_sampling_openai
1212 16285it [04:08, 11.14it/s]
    diversity_sampling_openai
1213 17623it [04:13, 270.28it/s]
    diversity_sampling_openai
1214 diversity_sampling_openai
1215 diversity_sampling_openai
1216 17642it [04:29, 5.11it/s]
    diversity_sampling_openai
1217 18961it [04:31, 596.91it/s]
    diversity_sampling_openai
1218 diversity_sampling_openai
1219 diversity_sampling_openai
1220 diversity_sampling_openai
1221 20343it [04:47, 658.10it/s]
    diversity_sampling_openai
1222 diversity_sampling_openai
1223 diversity_sampling_openai
1224 diversity_sampling_openai
1225 21695it [05:02, 645.06it/s]
    diversity_sampling_openai
1226 diversity_sampling_openai
1227 diversity_sampling_openai
1228 21713it [05:16, 14.07it/s]
    diversity_sampling_openai
1229 23010it [05:18, 643.03it/s]
    diversity_sampling_openai
1230 diversity_sampling_openai
1231 diversity_sampling_openai
1232 diversity_sampling_openai
1233 24401it [05:33, 571.64it/s]
    diversity_sampling_openai
1234 diversity_sampling_openai
1235 diversity_sampling_openai
1236 24427it [05:48, 11.74it/s]
    diversity_sampling_openai
1237 25766it [05:50, 463.75it/s]
    diversity_sampling_openai
1238 diversity_sampling_openai
1239 diversity_sampling_openai
1240 diversity_sampling_openai
```

```
1241 27126it [06:07, 409.59it/s]
    diversity_sampling_openai
1242 diversity_sampling_openai
1243 epoch number 19
1244 diversity_sampling_openai
1245 diversity_sampling_openai
1246 diversity_sampling_openai
1247 diversity_sampling_openai
1248 27140it [06:26, 409.59it/s]global t is : 7
1249 test epoch results {'accuracy': 0.5980392156862745
, 'f1_weighted': 0.5803314910032362, 'f1_macro': 0.
5181145191071922, 'f1': array([0.86582278, 0.
35087719, 0.55555556, 0.36571429, 0.58907363,
1250          0.13913043, 0.60550459, 0.61020036, 0.
58115183]), 'precision': array([0.855      , 0.5
, 0.6097561 , 0.62745098, 0.59903382,
1251          0.30769231, 0.64705882, 0.51857585, 0.
59677419]), 'recall': array([0.87692308, 0.27027027
, 0.51020408, 0.25806452, 0.57943925,
1252          0.08988764, 0.56896552, 0.74115044, 0.
56632653])}
1253 val epoch results {'accuracy': 0.6135986733001658
, 'f1_weighted': 0.5962694180238131, 'f1_macro': 0.
5178715283666809, 'f1': array([0.884273      , 0.
33333333, 0.42857143, 0.4025974 , 0.61585366,
1254          0.1686747 , 0.57142857, 0.61882353, 0.
63728814]), 'precision': array([0.86127168, 0.75
, 0.45454545, 0.56363636, 0.62345679,
1255          0.4375      , 0.61157025, 0.526      , 0.
68115942]), 'recall': array([0.90853659, 0.21428571
, 0.40540541, 0.31313131, 0.60843373,
1256          0.10447761, 0.53623188, 0.75142857, 0.
59872611])}
1257 27140it [06:26, 70.19it/s]
1258 diversity new embedding 2359
1259 diversity len embedding 2359
1260 MLP_1h(
1261     (l1): Linear(in_features=48, out_features=100,
bias=True)
1262     (l2): Linear(in_features=100, out_features=9,
bias=True)
```

1263 )  
1264 [1761, 2087, 1475, 1870, 885, 1751, 583, 1156, 603  
, 1271, 1391, 570, 170, 2080, 998, 1737, 1009, 749  
, 997, 1368, 2075, 125, 1968, 1398, 175, 212, 2224  
, 2064, 378, 600, 881, 1045, 553, 2221, 1659, 961,  
711, 1577, 1836, 564, 576, 2302, 829, 1783, 719,  
1668, 16, 486, 1021, 1225, 487, 944, 752, 925, 1397  
, 1342, 987, 2264, 2337, 1372, 2291, 1066, 184,  
2292, 108, 1586, 838, 75, 219, 201, 163, 1587, 121  
, 2348, 884, 308, 2253, 1877, 2136, 1472, 1525,  
1563, 13, 728, 1955, 1363, 646, 977, 1028, 821, 268  
, 35, 1395, 218, 948, 2257, 128, 1873, 2274, 578,  
464, 1700, 365, 388, 1990, 2234, 1282, 1062, 183,  
642, 1193, 2352, 1060, 714, 2056, 1419, 1071, 270,  
769, 899, 1628, 1413, 341, 65, 1582, 1429, 19, 996  
, 2275, 785, 623, 1369, 1327, 1565, 1913, 2090,  
2017, 320, 609, 251, 1560, 1507, 1883, 650, 72, 754  
, 236, 179, 323, 284, 1215, 880, 1833, 2313, 1010,  
528, 572, 348, 234, 1196, 739, 1013, 918, 1481,  
1875, 164, 28, 1026, 639, 1936, 1452, 1135, 1090,  
696, 84, 262, 974, 1209, 955, 2303, 439, 765, 322,  
841, 845, 2293, 1956, 708, 1803, 1173, 2190, 38,  
1501, 336, 2044, 1744, 631, 185, 2019, 644, 2071,  
1309, 116, 30, 1728, 1153, 337, 1771, 1618, 533,  
2110, 207, 802, 2285, 291, 2158, 10, 959, 1175, 763  
, 1722, 1355, 1792, 1816, 1409, 1034, 1538, 647,  
1037, 484, 2254, 2013, 423, 1755, 1250, 915, 1903,  
1685, 839, 1539, 1958, 1970, 1631, 2088, 1107, 1613  
, 1977, 1159, 394, 2282, 1115, 1124, 2045, 758,  
1575, 1174, 681, 1201, 2165, 15, 920, 1492, 424,  
1478, 2296, 1041, 192, 900, 305, 2118, 14, 129,  
2130, 2182, 1495, 2204, 1032, 1466, 2091, 50, 2265  
, 376, 511, 2199, 1313, 1141, 1580, 837, 1275, 1531  
, 742, 1770, 1396, 776, 1971, 2215, 1068, 1459,  
1098, 1235, 817, 591, 2125, 39, 781, 615, 1704,  
2197, 2070, 766, 1334, 34, 1320, 1023, 1953, 1094,  
2306, 223, 1125, 255, 2287, 1417, 1027, 947, 95,  
1177, 1491, 985, 2249, 832, 2037, 1218, 1121, 1137  
, 1614, 232, 2141, 1059, 1374, 2129, 2103, 1619,  
590, 2192, 1428, 410, 1272, 545, 1731, 1180, 2243,  
2195, 2353, 864, 1878, 1896, 1190, 92, 210, 2334,

1264 1420, 1567, 474, 147, 1784, 1670, 517, 786, 2295,  
2031, 1691, 1205, 1611, 197, 1321, 1463, 846, 1330  
, 1677, 849, 2351, 1488, 1743, 1978, 857, 796, 949  
, 1701, 1095, 1590, 2052, 2112, 2099, 620, 1217,  
1674, 1104, 2119, 2180, 1663, 2329, 1243, 237, 685  
, 1585, 908, 2145, 556, 1902, 1242, 1991, 1129,  
1553, 514, 436, 1088, 2051, 463, 901, 1349, 2171,  
588, 953, 297, 1182, 830, 2271, 428, 331, 1051,  
1219, 2240, 799, 173, 141, 555, 557, 1673, 1423, 64  
, 1181, 1373, 485, 783, 71, 1885, 978, 822, 1581,  
1089, 1723, 1, 250, 1666, 1049, 499, 1304, 2227,  
671, 1929, 640, 693, 814, 2211, 1739, 156, 958, 0,  
1621, 356, 2147]  
1265 Label: 7  
1266 generate a string that only contains the words in  
the text that are only semantically equal to other  
useful information:  
1267  
1268 Sleep-Tracking Data Shows Who Was Jolted Awake By  
The Napa Earthquake: Data from Jawbone Up users  
charts the m... http USA » NEWS » http HOT NEWS 550  
earthquake Northern California rattled by  
magnitude-6.0 earthquake. Dozens... Terrifying  
Animation Shows How Chile's Tsunami Took Over the  
Entire Pacific Ocean: [HTML1... http Tsunami travel  
time map (if a tsunami was generated). via Chile  
Peru http Napa earthquake losses could top \$1  
billion: http http .  
1269  
1270  
1271 AI Explanation: "Dozens Terrifying Animation Shows  
Pacific Ocean Napa Earthquake Losses Could Top  
Billion"  
1272 Downloading and preparing dataset csv/default to C  
:/Users/admin/.cache/huggingface/datasets/csv/  
default-f2a941eb9476e828/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8  
53bcc7ec513e1...  
1273 Downloading data files: 100%|██████████| 1/1 [00:00  
<, ?it/s]  
1274 Extracting data files: 100%|██████████| 1/1 [00:00<

```
1274 00:00, 171.43it/s]
1275 100%|██████████| 1/1 [00:00<00:00, 109.38it/s]
1276 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
f2a941eb9476e828/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1. Subsequent calls will reuse this
data.
1277 Downloading data files: 100%|██████████| 1/1 [00:00
<00:00, 1061.58it/s]
1278 Downloading and preparing dataset csv/default to C
:/Users/admin/.cache/huggingface/datasets/csv/
default-7c69f4c35d03f563/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1...
1279 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 168.41it/s]
1280 100%|██████████| 1/1 [00:00<00:00, 62.06it/s]
1281 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
7c69f4c35d03f563/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1. Subsequent calls will reuse this
data.
1282 Downloading data files: 100%|██████████| 1/1 [00:00
<, ?it/s]
1283 Downloading and preparing dataset csv/default to C
:/Users/admin/.cache/huggingface/datasets/csv/
default-7c7e93f4ed05e02c/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1...
1284 Extracting data files: 100%|██████████| 1/1 [00:00<
00:00, 160.33it/s]
1285 100%|██████████| 1/1 [00:00<00:00, 101.80it/s]
1286 Dataset csv downloaded and prepared to C:/Users/
admin/.cache/huggingface/datasets/csv/default-
7c7e93f4ed05e02c/0.0.0/
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
53bcc7ec513e1. Subsequent calls will reuse this
data.
1287 Downloading data files: 100%|██████████| 1/1 [00:00
```

```
1287 <?, ?it/s]
1288 Extracting data files: 100%|██████████| 1/1 [00:00<
    00:00, 155.78it/s]
1289  0% | 0/1 [00:00<?, ?it/s]Downloading
    and preparing dataset csv/default to C:/Users/admin
    /.cache/huggingface/datasets/csv/default-
    ef33c5a0cf5a35d3/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
    53bcc7ec513e1...
1290 Dataset csv downloaded and prepared to C:/Users/
    admin/.cache/huggingface/datasets/csv/default-
    ef33c5a0cf5a35d3/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
    53bcc7ec513e1. Subsequent calls will reuse this
    data.
1291 100%|██████████| 1/1 [00:00<00:00, 122.48it/s]
1292 t is : 8
1293 addlen is: 4000
1294 noexp_df.shape[0] last 1888
1295 unexplained dataset length is:
1296 1888
1297 embedding begin
1298 trian and val dataset length: 212670
1299 diversity_sampling_openai
1300 diversity_sampling_openai
1301 diversity_sampling_openai
1302 diversity_sampling_openai
1303 diversity_sampling_openai
1304 diversity_sampling_openai
1305 diversity_sampling_openai
1306 diversity_sampling_openai
1307 1251.0
1308 ./embeddings/NEW_bertie_embeddings_textattack/bert-
    base-uncased-MNLI_subset_1.pt
1309 Found cached dataset csv (C:/Users/admin/.cache/
    huggingface/datasets/csv/default-494a068ef8ae599d/0
    .0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
    53bcc7ec513e1)
1310 100%|██████████| 1/1 [00:00<00:00, 71.67it/s]
1311 test len 27285
```

```
1312 Downloading data files: 100%|██████████| 1/1 [00:00
    <?, ?it/s]
1313 Extracting data files: 100%|██████████| 1/1 [00:00<
    00:00, 154.88it/s]
1314 Generating train split: 0 examples [00:00, ?
    examples/s]Downloading and preparing dataset csv/
    default to C:/Users/admin/.cache/huggingface/
    datasets/csv/default-fe84b7cafcc23909/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
    53bcc7ec513e1...
1315 0%|          | 0/1 [00:00<?, ?it/s]Dataset csv
    downloaded and prepared to C:/Users/admin/.cache/
    huggingface/datasets/csv/default-fe84b7cafcc23909/0
    .0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
    53bcc7ec513e1. Subsequent calls will reuse this
    data.
1316 100%|██████████| 1/1 [00:00<00:00, 27.87it/s]
1317 preprocess_samples begin ...
1318 uncertainty shape: 1530
1319 ./test_embeddings/NEW_bertie_embeddings_textattack/
    bert-base-uncased-MNLI_subset_1.pt
1320 classify begin
1321 12564
1322 0%|          | 0/20 [00:00<?, ?it/s]
    diversity_sampling_openai
1323 diversity_sampling_openai
1324 1367it [00:17, 732.85it/s]diversity_sampling_openai
1325 diversity_sampling_openai
1326 diversity_sampling_openai
1327 diversity_sampling_openai
1328 2783it [00:32, 707.41it/s]diversity_sampling_openai
1329 diversity_sampling_openai
1330 diversity_sampling_openai
1331 2817it [00:45, 17.35it/s] diversity_sampling_openai
1332 4134it [00:47, 690.91it/s]diversity_sampling_openai
1333 diversity_sampling_openai
1334 diversity_sampling_openai
1335 diversity_sampling_openai
1336 5615it [01:02, 769.12it/s]diversity_sampling_openai
1337 diversity_sampling_openai
```

```
1338 diversity_sampling_openai
1339 5633it [01:15, 16.77it/s] diversity_sampling_openai
1340 6998it [01:16, 749.51it/s] diversity_sampling_openai
1341 diversity_sampling_openai
1342 diversity_sampling_openai
1343 diversity_sampling_openai
1344 8399it [01:32, 574.75it/s] diversity_sampling_openai
1345 diversity_sampling_openai
1346 diversity_sampling_openai
1347 diversity_sampling_openai
1348 9793it [01:48, 565.64it/s] diversity_sampling_openai
1349 diversity_sampling_openai
1350 diversity_sampling_openai
1351 diversity_sampling_openai
1352 11260it [02:04, 588.89it/s]
    diversity_sampling_openai
1353 diversity_sampling_openai
1354 diversity_sampling_openai
1355 diversity_sampling_openai
1356 12634it [02:19, 569.59it/s]
    diversity_sampling_openai
1357 diversity_sampling_openai
1358 diversity_sampling_openai
1359 12672it [02:32, 569.59it/s]
    diversity_sampling_openai
1360 14046it [02:35, 565.15it/s]
    diversity_sampling_openai
1361 diversity_sampling_openai
1362 diversity_sampling_openai
1363 diversity_sampling_openai
1364 15486it [02:50, 718.87it/s]
    diversity_sampling_openai
1365 diversity_sampling_openai
1366 diversity_sampling_openai
1367 15488it [03:02, 718.87it/s]
    diversity_sampling_openai
1368 16882it [03:06, 687.95it/s]
    diversity_sampling_openai
1369 diversity_sampling_openai
1370 diversity_sampling_openai
1371 diversity_sampling_openai
```

```
1372 18262it [03:21, 711.98it/s]
    diversity_sampling_openai
1373 diversity_sampling_openai
1374 diversity_sampling_openai
1375 18305it [03:35, 17.04it/s]
    diversity_sampling_openai
1376 19707it [03:37, 541.15it/s]
    diversity_sampling_openai
1377 diversity_sampling_openai
1378 diversity_sampling_openai
1379 diversity_sampling_openai
1380 21068it [03:53, 563.86it/s]
    diversity_sampling_openai
1381 diversity_sampling_openai
1382 diversity_sampling_openai
1383 diversity_sampling_openai
1384 22474it [04:08, 575.20it/s]
    diversity_sampling_openai
1385 diversity_sampling_openai
1386 diversity_sampling_openai
1387 22528it [04:22, 575.20it/s]
    diversity_sampling_openai
1388 23918it [04:24, 490.28it/s]
    diversity_sampling_openai
1389 diversity_sampling_openai
1390 diversity_sampling_openai
1391 diversity_sampling_openai
1392 25298it [04:40, 558.30it/s]
    diversity_sampling_openai
1393 diversity_sampling_openai
1394 diversity_sampling_openai
1395 25345it [04:55, 13.34it/s]
    diversity_sampling_openai
1396 26702it [04:56, 723.16it/s]
    diversity_sampling_openai
1397 diversity_sampling_openai
1398 diversity_sampling_openai
1399 diversity_sampling_openai
1400 28096it [05:11, 733.59it/s]
    diversity_sampling_openai
1401 diversity_sampling_openai
```

```
1402 epoch number 19
1403 diversity_sampling_openai
1404 diversity_sampling_openai
1405 28160it [05:22, 733.59it/s]
    diversity_sampling_openai
1406 diversity_sampling_openai
1407 global t is : 8
1408 test epoch results {'accuracy': 0.6058823529411764
, 'f1_weighted': 0.5904350247389664, 'f1_macro': 0.
516525162355184, 'f1': array([0.86284289, 0.
26229508, 0.5 , 0.36158192, 0.60496614,
1409 0.22900763, 0.60895522, 0.63018868, 0.
58888889]), 'precision': array([0.83980583, 0.
33333333, 0.53488372, 0.60377358, 0.58515284,
1410 0.35714286, 0.63354037, 0.54934211, 0.
64634146]), 'recall': array([0.88717949, 0.21621622
, 0.46938776, 0.25806452, 0.62616822,
1411 0.16853933, 0.5862069 , 0.73893805, 0.
54081633])}
1412 val epoch results {'accuracy': 0.6083133493205436
, 'f1_weighted': 0.5930280191248222, 'f1_macro': 0.
5236101254022297, 'f1': array([0.85714286, 0.
27027027, 0.52777778, 0.51190476, 0.60655738,
1413 0.16666667, 0.58088235, 0.61501788, 0.
57627119]), 'precision': array([0.81382979, 0.625
, 0.57575758, 0.66153846, 0.57216495,
1414 0.32 , 0.6124031 , 0.54315789, 0.
63432836]), 'recall': array([0.90532544, 0.17241379
, 0.48717949, 0.41747573, 0.64534884,
1415 0.11267606, 0.55244755, 0.70879121, 0.
52795031])}
1416 diversity new embedding 1888
1417 diversity len embedding 1888
1418 MLP_1h(
1419     (l1): Linear(in_features=51, out_features=100,
        bias=True)
1420     (l2): Linear(in_features=100, out_features=9,
        bias=True)
1421 )
1422 28160it [05:28, 85.70it/s]
1423 [578, 1748, 1008, 1310, 1869, 802, 358, 897, 1308,
```

|      |   |
|------|---|
| 1423 | 1031, 998, 1857, 227, 395, 263, 1416, 710, 787, 1850, 1330, 83, 1510, 582, 273, 931, 285, 1313, 1783, 777, 768, 472, 68, 1773, 1795, 439, 1063, 1372, 447, 1778, 946, 40, 1409, 515, 1558, 364, 633, 624, 196, 1091, 1583, 1301, 1841, 1133, 1826, 1296, 690, 1828, 1870, 431, 102, 927, 1507, 646, 1318, 186, 1189, 1290, 1781, 208, 906, 1049, 541, 28, 645, 1289, 379, 586, 1214, 427, 162, 374, 687, 704, 420, 35, 580, 1227, 310, 1574, 898, 1401, 1383, 932, 1655, 607, 1705, 150, 1027, 501, 911, 1636, 1349, 602, 1270, 1482, 222, 1457, 850, 122, 1494, 764, 1462, 1496, 556, 910, 334, 1665, 1370, 727, 555, 1134, 1648, 1093, 1048, 1014, 945, 1005, 1662, 542, 293, 193, 99, 1395, 511, 1261, 1020, 1807, 756, 598, 1703, 1366, 157, 1814, 884, 1572, 746, 641, 187, 311, 860, 676, 1158, 896, 1673, 81, 1853, 1228, 276, 1024, 697, 1678, 1737, 1042, 1069, 706, 493, 1272, 197, 959, 913, 565, 1791, 1078, 707, 394, 1193, 494, 1044, 1025, 1368, 1697, 342, 214, 115, 836, 507, 992, 51, 985, 1758, 549, 1535, 1127, 14, 686, 1403, 1087, 1075, 1052, 845, 1530, 1516, 478, 1190, 934, 905, 879, 865, 274, 1860, 291, 368, 1146, 1168, 1298, 215, 135, 1883, 1466, 1361, 1464, 1491, 660, 258, 134, 1082, 1321, 242, 1179, 97, 320, 1336, 1249, 863, 735, 57, 325, 978, 1753, 1514, 1863, 821, 647, 1639, 463, 1374, 675, 486, 496, 599, 1620, 1036, 767, 952, 417, 89, 490, 920, 1563, 238, 816, 390, 1204, 1536, 1858, 557, 506, 221, 1881, 790, 161, 182, 1624, 4, 1539, 1424, 939, 1367, 1445, 303, 871, 1335, 1355, 524, 243, 517, 1411, 412, 306, 1670, 145, 1804, 592, 136, 1398, 872, 25, 65, 758, 792, 1066, 1456, 804, 949, 1793, 1731, 1142, 200, 1730, 708, 459, 497, 1459, 378, 1030, 393, 1385, 654, 1284, 268, 1794, 1253, 1012, 1109, 1469, 1570, 755, 1543, 665, 858, 1637, 1410, 1835, 838, 1738, 1843, 1431, 886, 445, 1677, 1010, 1266, 219, 414, 1371, 376, 363, 1297, 937, 782, 18, 1763, 1736, 370, 1382, 1546, 1805, 331, 1844, 1269, 785, 1413, 1081, 127, 1875, 1350, 1513, 7, 1102, 1169, 914, 1521, 1278, 1597, 1554, 235, 1420] |
| 1424 | Label: 8  |

1425 generate a string that only contains the words in  
the text that are only semantically equal to not  
related or irrelevant:

1426

1427 Nessuno tsunami in Cile! Largest swimming pool in  
the world at the San Alfonso del Mar Resort Chile  
http Obama with his family. Pakistani family  
killed by Obama's drones. U S Crimes U S Pakistan  
http I miss waking up to long sweet good morning  
text messages neutral faceweary facered  
hearthundred points &gt;&gt;&gt; Wanted a cyclone  
ice block and then realised that my brother ate the  
last one today why B Noooooooooo "What About  
The Wine?!" re Napa, California, M6.1 quake, via  
http.

1428

1429

1430 AI Explanation: "Nessuno tsunami in Cile!" "U S  
Crimes U S Pakistan" "I miss waking up to long

1431 Downloading data files: 100%|██████████| 1/1 [00:00  
<, ?it/s]

1432 Extracting data files: 100%|██████████| 1/1 [00:00<  
00:00, 167.05it/s]

1433 Downloading and preparing dataset csv/default to C  
:/Users/admin/.cache/huggingface/datasets/csv/  
default-32a163c36cf0e790/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8  
53bcc7ec513e1...

1434 Dataset csv downloaded and prepared to C:/Users/  
admin/.cache/huggingface/datasets/csv/default-  
32a163c36cf0e790/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8  
53bcc7ec513e1. Subsequent calls will reuse this  
data.

1435 100%|██████████| 1/1 [00:00<00:00, 77.19it/s]

1436 Downloading data files: 100%|██████████| 1/1 [00:00  
<, ?it/s]

1437 Downloading and preparing dataset csv/default to C  
:/Users/admin/.cache/huggingface/datasets/csv/  
default-4d9ba21811f05aa2/0.0.0/  
6954658bab30a358235fa864b05cf819af0e179325c740e4bc8

```
1437 53bcc7ec513e1...
1438 Extracting data files: 100%|██████████| 1/1 [00:00<
    00:00, 165.38it/s]
1439 100%|██████████| 1/1 [00:00<00:00, 50.15it/s]
1440 Dataset csv downloaded and prepared to C:/Users/
    admin/.cache/huggingface/datasets/csv/default-
    4d9ba21811f05aa2/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
    53bcc7ec513e1. Subsequent calls will reuse this
    data.
1441 Downloading data files: 100%|██████████| 1/1 [00:00
    <, ?it/s]
1442 Extracting data files: 0%|          | 0/1 [00:00
    <, ?it/s]Downloading and preparing dataset csv/
    default to C:/Users/admin/.cache/huggingface/
    datasets/csv/default-e2347f3bdf07fd5f/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
    53bcc7ec513e1...
1443 Extracting data files: 100%|██████████| 1/1 [00:00<
    00:00, 37.16it/s]
1444 100%|██████████| 1/1 [00:00<00:00, 76.86it/s]
1445 Dataset csv downloaded and prepared to C:/Users/
    admin/.cache/huggingface/datasets/csv/default-
    e2347f3bdf07fd5f/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
    53bcc7ec513e1. Subsequent calls will reuse this
    data.
1446 Downloading data files: 100%|██████████| 1/1 [00:00
    <, ?it/s]
1447 Downloading and preparing dataset csv/default to C
    :/Users/admin/.cache/huggingface/datasets/csv/
    default-ffb623b82b861abb/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
    53bcc7ec513e1...
1448 Extracting data files: 100%|██████████| 1/1 [00:00<
    00:00, 125.41it/s]
1449 100%|██████████| 1/1 [00:00<00:00, 111.43it/s]
1450 Dataset csv downloaded and prepared to C:/Users/
    admin/.cache/huggingface/datasets/csv/default-
    ffb623b82b861abb/0.0.0/
    6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
```

```
1450 53bcc7ec513e1. Subsequent calls will reuse this
      data.
1451 t is : 9
1452 addlen is: 4500
1453 noexp_df.shape[0] last 1511
1454 unexplained dataset length is:
1455 1511
1456 embedding begin
1457 trian and val dataset length: 231660
1458 diversity_sampling_openai
1459 diversity_sampling_openai
1460 diversity_sampling_openai
1461 diversity_sampling_openai
1462 diversity_sampling_openai
1463 diversity_sampling_openai
1464 diversity_sampling_openai
1465 diversity_sampling_openai
1466 1287.0
1467 ./embeddings/NEW_bertie_embeddings_textattack/bert-
      base-uncased-MNLI_subset_1.pt
1468 Found cached dataset csv (C:/Users/admin/.cache/
      huggingface/datasets/csv/default-494a068ef8ae599d/0
      .0.0/
      6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
      53bcc7ec513e1)
1469 100%|██████████| 1/1 [00:00<00:00, 55.73it/s]
1470 test len 28890
1471 Downloading data files: 100%|██████████| 1/1 [00:00
      <, ?it/s]
1472 Extracting data files: 100%|██████████| 1/1 [00:00<
      00:00, 125.73it/s]
1473 Generating train split: 0 examples [00:00, ?
      examples/s]Downloading and preparing dataset csv/
      default to C:/Users/admin/.cache/huggingface/
      datasets/csv/default-37bd5c9152f24e61/0.0.0/
      6954658bab30a358235fa864b05cf819af0e179325c740e4bc8
      53bcc7ec513e1...
1474 0%|          | 0/1 [00:00<?, ?it/s]Dataset csv
      downloaded and prepared to C:/Users/admin/.cache/
      huggingface/datasets/csv/default-37bd5c9152f24e61/0
      .0.0/
```

```
1474      6954658bab30a358235fa864b05cf819af0e179325c740e4bc8  
53bcc7ec513e1. Subsequent calls will reuse this  
data.  
1475 100% [██████] 1/1 [00:00<00:00, 28.67it/s]  
1476 preprocess_samples begin ...  
1477 uncertainty shape: 1530  
1478 ./test_embeddings/NEW_bertie_embeddings_textattack/  
bert-base-uncased-MNLI_subset_1.pt  
1479 classify begin  
1480 12941  
1481 0% | 0/20 [00:00<?, ?it/s]  
    diversity_sampling_openai  
1482 diversity_sampling_openai  
1483 1405it [00:19, 658.45it/s]diversity_sampling_openai  
1484 diversity_sampling_openai  
1485 diversity_sampling_openai  
1486 diversity_sampling_openai  
1487 2872it [00:34, 646.32it/s]diversity_sampling_openai  
1488 diversity_sampling_openai  
1489 diversity_sampling_openai  
1490 2897it [00:48, 14.68it/s] diversity_sampling_openai  
1491 4331it [00:52, 457.22it/s]diversity_sampling_openai  
1492 diversity_sampling_openai  
1493 diversity_sampling_openai  
1494 4345it [01:06, 9.06it/s] diversity_sampling_openai  
1495 5755it [01:08, 537.66it/s]diversity_sampling_openai  
1496 diversity_sampling_openai  
1497 diversity_sampling_openai  
1498 diversity_sampling_openai  
1499 7225it [01:25, 467.71it/s]diversity_sampling_openai  
1500 diversity_sampling_openai  
1501 diversity_sampling_openai  
1502 diversity_sampling_openai  
1503 8668it [01:42, 525.67it/s]diversity_sampling_openai  
1504 diversity_sampling_openai  
1505 diversity_sampling_openai  
1506 8689it [01:56, 10.34it/s] diversity_sampling_openai  
1507 10085it [01:58, 547.70it/s]  
    diversity_sampling_openai  
1508 diversity_sampling_openai
```

```
1509 diversity_sampling_openai
1510 diversity_sampling_openai
1511 11547it [02:15, 487.47it/s]
    diversity_sampling_openai
1512 diversity_sampling_openai
1513 diversity_sampling_openai
1514 diversity_sampling_openai
1515 12996it [02:31, 720.35it/s]
    diversity_sampling_openai
1516 diversity_sampling_openai
1517 diversity_sampling_openai
1518 diversity_sampling_openai
1519 14435it [02:46, 619.34it/s]
    diversity_sampling_openai
1520 diversity_sampling_openai
1521 diversity_sampling_openai
1522 diversity_sampling_openai
1523 15898it [03:02, 649.15it/s]
    diversity_sampling_openai
1524 diversity_sampling_openai
1525 diversity_sampling_openai
1526 15929it [03:16, 15.29it/s]
    diversity_sampling_openai
1527 17346it [03:18, 630.46it/s]
    diversity_sampling_openai
1528 diversity_sampling_openai
1529 diversity_sampling_openai
1530 diversity_sampling_openai
1531 18816it [03:33, 670.95it/s]
    diversity_sampling_openai
1532 diversity_sampling_openai
1533 diversity_sampling_openai
1534 18824it [03:45, 670.95it/s]
    diversity_sampling_openai
1535 20266it [03:51, 362.77it/s]
    diversity_sampling_openai
1536 diversity_sampling_openai
1537 diversity_sampling_openai
1538 diversity_sampling_openai
1539 21678it [04:07, 543.96it/s]
    diversity_sampling_openai
```

```
1540 diversity_sampling_openai
1541 diversity_sampling_openai
1542 diversity_sampling_openai
1543 23155it [04:24, 518.36it/s]
    diversity_sampling_openai
1544 diversity_sampling_openai
1545 diversity_sampling_openai
1546 23169it [04:38, 10.23it/s]
    diversity_sampling_openai
1547 24582it [04:40, 448.79it/s]
    diversity_sampling_openai
1548 diversity_sampling_openai
1549 diversity_sampling_openai
1550 diversity_sampling_openai
1551 26034it [04:57, 526.13it/s]
    diversity_sampling_openai
1552 diversity_sampling_openai
1553 diversity_sampling_openai
1554 diversity_sampling_openai
1555 27484it [05:14, 483.99it/s]
    diversity_sampling_openai
1556 diversity_sampling_openai
1557 diversity_sampling_openai
1558 27513it [05:27, 10.42it/s]
    diversity_sampling_openai
1559 28894it [05:29, 662.75it/s]
    diversity_sampling_openai
1560 diversity_sampling_openai
1561 epoch number 19
1562 diversity_sampling_openai
1563 diversity_sampling_openai
1564 28960it [05:45, 662.75it/s]
    diversity_sampling_openai
1565 diversity_sampling_openai
1566 global t is : 9
1567 test epoch results {'accuracy': 0.6143790849673203
, 'f1_weighted': 0.6000679210175794, 'f1_macro': 0.
535269402234186, 'f1': array([0.865      , 0.
37288136, 0.48837209, 0.40659341, 0.61538462,
1568      0.23076923, 0.60493827, 0.63293468, 0.
60055096]), 'precision': array([0.84390244, 0.5
```

```

1568      , 0.56756757, 0.63793103, 0.61395349,
1569      0.36585366, 0.65333333, 0.54173228, 0.
       65269461]), 'recall': array([0.88717949, 0.2972973
       , 0.42857143, 0.2983871 , 0.61682243,
1570      0.16853933, 0.56321839, 0.76106195, 0.
       55612245])}
1571 val epoch results {'accuracy': 0.6293706293706294
       , 'f1_weighted': 0.6148421084686472, 'f1_macro': 0.
       5379625001540441, 'f1': array([0.86027397, 0.
       29268293, 0.39393939, 0.44578313, 0.63817664,
1572      0.33009709, 0.61538462, 0.63787375, 0.
       62745098]), 'precision': array([0.81770833, 0.
       54545455, 0.5      , 0.62711864, 0.64367816,
1573      0.56666667, 0.66666667, 0.54442344, 0.
       68571429]), 'recall': array([0.90751445, 0.2
       , 0.325      , 0.34579439, 0.63276836,
1574      0.23287671, 0.57142857, 0.77005348, 0.
       57831325])}
1575 28960it [05:46, 83.48it/s]
1576 diversity new embedding 1511
1577 diversity len embedding 1511
1578 MLP_1h(
1579   (l1): Linear(in_features=54, out_features=100,
     bias=True)
1580   (l2): Linear(in_features=100, out_features=9,
     bias=True)
1581 )
1582 [35, 587, 694, 1385, 1295, 693, 569, 255, 418, 582
       , 761, 1510, 1285, 249, 376, 136, 743, 244, 1124,
       1087, 695, 1296, 513, 15, 1289, 66, 8, 292, 952,
       919, 1495, 1250, 637, 51, 468, 610, 1458, 324, 830
       , 107, 596, 955, 669, 680, 344, 353, 430, 1255, 47
       , 1312, 1336, 721, 398, 665, 568, 123, 1272, 719,
       261, 959, 1196, 297, 620, 588, 1355, 1080, 75, 609
       , 278, 621, 966, 631, 1346, 660, 273, 327, 88, 359
       , 1502, 1234, 154, 814, 698, 1310, 613, 608, 1298,
       1405, 1411, 1396, 811, 1450, 453, 733, 699, 152,
       944, 130, 1045, 114, 1174, 799, 903, 871, 476, 506
       , 1472, 827, 109, 459, 1269, 950, 810, 542, 471,
       1437, 517, 907, 288, 793, 1433, 42, 253, 545, 1074
       , 848, 1417, 266, 183, 112, 1429, 726, 860, 898,

```

```

1582 1420, 1408, 839, 1058, 861, 1104, 1165, 436, 105,
      584, 118, 1481, 1133, 1021, 1392, 1359, 808, 300,
      110, 58, 1406, 511, 65, 562, 1379, 37, 463, 59, 193
      , 720, 1257, 1287, 443, 407, 337, 875, 817, 1044,
      280, 804, 1283, 1175, 1291, 742, 357, 994, 1187,
      934, 807, 393, 38, 175, 767, 1129, 1395, 474, 1144
      , 1452, 999, 1443, 1064, 106, 484, 854, 1247, 1292
      , 18, 1212, 1100, 540, 731, 194, 1033, 82, 5, 1014
      , 1308, 909, 320, 522, 753, 440, 638, 874, 911, 184
      , 850, 485, 1084, 691, 1338, 975, 390, 519, 71, 925
      , 1263, 696, 902, 186, 388, 749, 1008, 373, 1024, 2
      , 1106, 1268, 210, 73, 554, 1022, 537, 1207, 134,
      1262, 927, 1026, 1309, 399, 592, 1278, 135, 639,
      844, 229, 1383, 69, 772, 865, 1451, 625, 565, 28,
      1504, 342, 1389, 526, 735, 263, 1173, 264, 1085,
      1156, 20, 980, 672, 1151, 424, 1170, 1162, 1299,
      922, 1363, 1251, 493, 1381, 1177, 1500, 549, 1508,
      1438, 1178, 1213, 128, 77, 732, 1466]
1583 Label: 9
1584 Traceback (most recent call last):
1585   File "D:\Bristol course\Data Science Dissertation
      \ExpBERT lab and thesis\BERT-IE-main\BERT-IE-main\
      diversity_sampling_whole.py", line 1443, in <module>
1586     main()
1587   File "D:\Bristol course\Data Science Dissertation
      \ExpBERT lab and thesis\BERT-IE-main\BERT-IE-main\
      diversity_sampling_whole.py", line 1432, in main
1588     addOrDelete(sampled_indices,raw_dataset_noexp)
1589   File "D:\Bristol course\Data Science Dissertation
      \ExpBERT lab and thesis\BERT-IE-main\BERT-IE-main\
      diversity_sampling_whole.py", line 612, in
      addOrDelete
1590     exp_list = generate_explanations_AI(
      sampled_data)
1591   File "D:\Bristol course\Data Science Dissertation
      \ExpBERT lab and thesis\BERT-IE-main\BERT-IE-main\
      diversity_sampling_whole.py", line 582, in
      generate_explanations_AI
1592     x = "generate a string that only contains the
      words in the text that are only semantically equal

```

File - diversity\_sampling\_whole

```
1592 to "+textual_descriptions[t]+": "+"\\n\\n"
1593 IndexError: list index out of range
1594
1595 Process finished with exit code 1
1596
```