**Assignment‌ ‌No:‌ ‌4‌‌**

**‌Prajakta‌ ‌Mane(BE/A/B1821016)‌**

**Title**: Yelp reviews polarity

**Objectives:**

1. To handle given data for Text Classification

2. To analyze Text using python programming language.

**Software Requirement:**

**Windows /Linux**

**Theory:**

The Yelp reviews polarity dataset is constructed by considering stars 1 and 2 negative, and 3 and 4 positive. For each polarity 280,000 training samples and 19,000 testing samples are take randomly. In total there are 560,000 trainig samples and 38,000 testing samples. Negative polarity is class 1, and positive class 2.

Definition:

―Information processing model that is inspired by the way biological nervous system (i.e. the brain) process information, is called Neural Network.

Neural Network has the ability to learn by examples. It is not designed to perform fix /, rather task which need thinking (e.g. Predictions).

ANN is composed of large number of highly interconnected processing elements(neurons)working in unison to solve problems. It mimic human brain.

With advanced in deep learning, you can now visualize the entire deep learning process or just the Convolutional Neural Network you’ve built.

We are going to build simple neural network using Keras and then use ANN visualizer to visualize our neural network.

CODE :

from tensorflow.keras.models import Sequential

from tensorflow.keras.layers import Dense, Embedding, GRU, LSTM, Dropout

from tensorflow.keras import utils

from tensorflow.keras.preprocessing.sequence import pad\_sequences

from tensorflow.keras.preprocessing.text import Tokenizer

from tensorflow.keras.callbacks import ModelCheckpoint, ReduceLROnPlateau

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from google.colab import drive

drive.mount('/content/drive')

Mounted at /content/drive

num\_words = 10000

max\_review\_len = 200

train = pd.read\_csv('/content/drive/MyDrive/A6/train.csv',

header=None,

names=['Label', 'Review'])

test = pd.read\_csv('/content/drive/MyDrive/A6/test.csv',

header=None,

names=['Label', 'Review'])

train

Label Review

0 1 Unfortunately, the frustration of being Dr. Go...

1 2 Been going to Dr. Goldberg for over 10 years. ...

2 1 I don't know what Dr. Goldberg was like before...

3 1 I'm writing this review to give you a heads up...

4 2 All the food is great here. But the best thing...

... ... ...

559995 2 Ryan was as good as everyone on yelp has claim...

559996 2 Professional \nFriendly\nOn time AND affordabl...

559997 1 Phone calls always go to voicemail and message...

559998 1 Looks like all of the good reviews have gone t...

559999 2 Ryan Rocks! I called him this morning for some...

[560000 rows x 2 columns]

y\_train, y\_test = train['Label'] - 1, test['Label'] - 1

reviews = train['Review']

reviews[:5]

0 Unfortunately, the frustration of being Dr. Go...

1 Been going to Dr. Goldberg for over 10 years. ...

2 I don't know what Dr. Goldberg was like before...

3 I'm writing this review to give you a heads up...

4 All the food is great here. But the best thing...

Name: Review, dtype: object

tokenizer = Tokenizer(num\_words=num\_words)

tokenizer.fit\_on\_texts(reviews)

tokenizer.word\_index

{'the': 1,

'and': 2,

'i': 3,

'to': 4,

'a': 5,

'was': 6,

'of': 7,

'it': 8,

'for': 9,

'in': 10,

'is': 11,

'n': 12,

'that': 13,

'my': 14,

'we': 15,

'this': 16,

'but': 17,

'with': 18,

'they': 19,

'you': 20,

'on': 21,

'not': 22,

'have': 23,

'had': 24,

'at': 25,

'were': 26,

'so': 27,

'are': 28,

'food': 29,

'be': 30,

'place': 31,

'me': 32,

'there': 33,

'good': 34,

'as': 35,

'out': 36,

'all': 37,

'like': 38,

'if': 39,

'just': 40,

'our': 41,

'very': 42,

'get': 43,

'one': 44,

'here': 45,

'time': 46,

'when': 47,

'up': 48,

'or': 49,

'from': 50,

'great': 51,

'service': 52,

'would': 53,

'back': 54,

'their': 55,

'about': 56,

'no': 57,

'go': 58,

'an': 59,

'what': 60,

'really': 61,

'he': 62,

"it's": 63,

'she': 64,

'which': 65,

'will': 66,

'some': 67,

'only': 68,

'been': 69,

'us': 70,

'your': 71,

'because': 72,

'more': 73,

'even': 74,

'can': 75,

'them': 76,

'by': 77,

"don't": 78,

'got': 79,

'other': 80,

'after': 81,

'do': 82,

'nthe': 83,

'also': 84,

'ni': 85,

'too': 86,

"didn't": 87,

'never': 88,

'ordered': 89,

'did': 90,

'well': 91,

'nice': 92,

"i'm": 93,

'over': 94,

'has': 95,

'order': 96,

'then': 97,

'went': 98,

'than': 99,

'people': 100,

'could': 101,

'came': 102,

'much': 103,

'restaurant': 104,

'first': 105,

'said': 106,

'her': 107,

'again': 108,

'off': 109,

'little': 110,

'best': 111,

"i've": 112,

'know': 113,

'better': 114,

'2': 115,

'going': 116,

'how': 117,

'who': 118,

'way': 119,

'two': 120,

'try': 121,

'love': 122,

'make': 123,

'come': 124,

'chicken': 125,

'pretty': 126,

'room': 127,

'always': 128,

'staff': 129,

'minutes': 130,

'menu': 131,

'vegas': 132,

'night': 133,

'any': 134,

'bar': 135,

'before': 136,

'experience': 137,

'made': 138,

'down': 139,

'want': 140,

'told': 141,

'think': 142,

'bad': 143,

'am': 144,

'say': 145,

'right': 146,

'day': 147,

'5': 148,

'asked': 149,

'still': 150,

'ever': 151,

'take': 152,

'around': 153,

'wait': 154,

'see': 155,

'table': 156,

'another': 157,

'eat': 158,

'new': 159,

'took': 160,

'friendly': 161,

'few': 162,

'give': 163,

'3': 164,

'now': 165,

'while': 166,

'since': 167,

"wasn't": 168,

'sure': 169,

'definitely': 170,

'pizza': 171,

'next': 172,

'being': 173,

'something': 174,

'times': 175,

'cheese': 176,

'his': 177,

'into': 178,

'where': 179,

'most': 180,

'sauce': 181,

'last': 182,

'everything': 183,

'thing': 184,

'find': 185,

'price': 186,

'delicious': 187,

'should': 188,

'many': 189,

'long': 190,

'meal': 191,

'wanted': 192,

'every': 193,

'drinks': 194,

'though': 195,

'salad': 196,

'small': 197,

'location': 198,

'lunch': 199,

'1': 200,

'10': 201,

'4': 202,

'lot': 203,

'check': 204,

'area': 205,

'amazing': 206,

'fresh': 207,

"can't": 208,

'nothing': 209,

'dinner': 210,

'bit': 211,

'its': 212,

'hotel': 213,

'left': 214,

'work': 215,

'same': 216,

'drink': 217,

'customer': 218,

'big': 219,

'tried': 220,

'hour': 221,

'side': 222,

'both': 223,

'worth': 224,

'looking': 225,

'away': 226,

'why': 227,

'hot': 228,

'him': 229,

'home': 230,

'need': 231,

'store': 232,

'enough': 233,

'server': 234,

'burger': 235,

'recommend': 236,

'stars': 237,

"you're": 238,

'taste': 239,

'called': 240,

'things': 241,

'quality': 242,

"that's": 243,

'ok': 244,

'happy': 245,

'different': 246,

'actually': 247,

'money': 248,

'feel': 249,

'anything': 250,

'free': 251,

'thought': 252,

'manager': 253,

'old': 254,

'fries': 255,

'prices': 256,

'these': 257,

'meat': 258,

'look': 259,

'line': 260,

'nwe': 261,

'else': 262,

'friends': 263,

'awesome': 264,

'ask': 265,

'found': 266,

'sushi': 267,

'decided': 268,

'looked': 269,

'top': 270,

'put': 271,

'once': 272,

'maybe': 273,

'years': 274,

'probably': 275,

'front': 276,

'however': 277,

'pay': 278,

'flavor': 279,

'half': 280,

'full': 281,

'done': 282,

'car': 283,

'sandwich': 284,

'visit': 285,

'clean': 286,

'getting': 287,

'review': 288,

'places': 289,

'call': 290,

'waitress': 291,

'through': 292,

'beer': 293,

'water': 294,

'friend': 295,

'show': 296,

'business': 297,

'least': 298,

'used': 299,

'20': 300,

'special': 301,

'stay': 302,

'favorite': 303,

'each': 304,

'bread': 305,

'finally': 306,

'sweet': 307,

'coffee': 308,

'steak': 309,

'star': 310,

'let': 311,

'rice': 312,

'breakfast': 313,

'kind': 314,

'super': 315,

'gave': 316,

'waiting': 317,

'30': 318,

'those': 319,

'tell': 320,

'tasted': 321,

'inside': 322,

'huge': 323,

'strip': 324,

'use': 325,

'almost': 326,

'open': 327,

'someone': 328,

'without': 329,

'later': 330,

'dish': 331,

'hard': 332,

'excellent': 333,

'quite': 334,

'fried': 335,

'reviews': 336,

'disappointed': 337,

'high': 338,

'having': 339,

'large': 340,

'far': 341,

'selection': 342,

'couple': 343,

"couldn't": 344,

'coming': 345,

'beef': 346,

'person': 347,

'care': 348,

'atmosphere': 349,

'buffet': 350,

'house': 351,

'tasty': 352,

'busy': 353,

'everyone': 354,

'fun': 355,

'served': 356,

'walked': 357,

'guy': 358,

'outside': 359,

'hours': 360,

'soup': 361,

'trying': 362,

'15': 363,

"i'd": 364,

'tables': 365,

'wrong': 366,

'end': 367,

'until': 368,

'husband': 369,

'town': 370,

'wine': 371,

"won't": 372,

'eating': 373,

'seemed': 374,

'whole': 375,

"i'll": 376,

'part': 377,

'three': 378,

'less': 379,

'items': 380,

'such': 381,

'waiter': 382,

'fish': 383,

'worst': 384,

'cold': 385,

'cooked': 386,

'walk': 387,

'perfect': 388,

'instead': 389,

'6': 390,

'family': 391,

'usually': 392,

'decent': 393,

'door': 394,

'must': 395,

'waited': 396,

'cream': 397,

"doesn't": 398,

'horrible': 399,

'during': 400,

'rude': 401,

'week': 402,

'brought': 403,

'point': 404,

'either': 405,

'does': 406,

'restaurants': 407,

'close': 408,

'cool': 409,

'needed': 410,

'felt': 411,

'second': 412,

'shrimp': 413,

'dishes': 414,

'extra': 415,

'several': 416,

'plate': 417,

'deal': 418,

'music': 419,

'year': 420,

'course': 421,

'started': 422,

'enjoy': 423,

'pork': 424,

'days': 425,

'help': 426,

'party': 427,

'oh': 428,

'arrived': 429,

"isn't": 430,

'shop': 431,

'saw': 432,

'ice': 433,

'spot': 434,

'reason': 435,

'customers': 436,

'may': 437,

'sat': 438,

'own': 439,

'yes': 440,

'leave': 441,

'especially': 442,

'name': 443,

'loved': 444,

'nmy': 445,

'phone': 446,

'fast': 447,

'owner': 448,

'dining': 449,

"wouldn't": 450,

'guess': 451,

'cheap': 452,

'job': 453,

'keep': 454,

'wife': 455,

'dessert': 456,

'return': 457,

'bring': 458,

'might': 459,

'problem': 460,

'comes': 461,

'terrible': 462,

'able': 463,

'yelp': 464,

'seated': 465,

'fine': 466,

'sit': 467,

'group': 468,

'paid': 469,

'dry': 470,

'spicy': 471,

'makes': 472,

'enjoyed': 473,

'already': 474,

'kids': 475,

'anyone': 476,

'red': 477,

"there's": 478,

'8': 479,

'stop': 480,

'parking': 481,

'las': 482,

'fact': 483,

'yet': 484,

'expect': 485,

'bill': 486,

'plus': 487,

'okay': 488,

'floor': 489,

'real': 490,

'today': 491,

'quick': 492,

'myself': 493,

'extremely': 494,

'absolutely': 495,

'7': 496,

'start': 497,

'chocolate': 498,

'club': 499,

'drive': 500,

'cut': 501,

'expensive': 502,

'chips': 503,

'late': 504,

'overall': 505,

'ate': 506,

'pool': 507,

'stuff': 508,

'slow': 509,

'ago': 510,

'guys': 511,

'50': 512,

'hair': 513,

'working': 514,

'charge': 515,

'card': 516,

'rooms': 517,

'offer': 518,

'seems': 519,

'doing': 520,

'buy': 521,

'options': 522,

'live': 523,

'ended': 524,

'tea': 525,

'fan': 526,

'roll': 527,

'nthis': 528,

'although': 529,

'rather': 530,

'trip': 531,

'local': 532,

'street': 533,

'sitting': 534,

'size': 535,

'making': 536,

'kitchen': 537,

'mexican': 538,

'style': 539,

'given': 540,

'wonderful': 541,

'bland': 542,

'tacos': 543,

'pick': 544,

'wish': 545,

'past': 546,

'seen': 547,

'rolls': 548,

'helpful': 549,

'bacon': 550,

'liked': 551,

'completely': 552,

'counter': 553,

'seem': 554,

'fantastic': 555,

'casino': 556,

'short': 557,

'etc': 558,

'decor': 559,

'girl': 560,

'regular': 561,

'looks': 562,

'run': 563,

'poor': 564,

'under': 565,

'believe': 566,

'mind': 567,

"they're": 568,

'company': 569,

'dog': 570,

'thai': 571,

'morning': 572,

'highly': 573,

'amount': 574,

'entire': 575,

'change': 576,

'non': 577,

'tip': 578,

'received': 579,

'lots': 580,

'ready': 581,

'employees': 582,

'average': 583,

'white': 584,

'lady': 585,

'green': 586,

'twice': 587,

'impressed': 588,

'gone': 589,

'man': 590,

'00': 591,

'kept': 592,

'heard': 593,

'set': 594,

'soon': 595,

'empty': 596,

'cake': 597,

'crab': 598,

'between': 599,

'taking': 600,

'bbq': 601,

'saturday': 602,

'seriously': 603,

'behind': 604,

'glass': 605,

'list': 606,

'understand': 607,

'serve': 608,

'portions': 609,

'early': 610,

'quickly': 611,

'offered': 612,

'light': 613,

'desk': 614,

'rest': 615,

'wings': 616,

'four': 617,

'warm': 618,

'burgers': 619,

'beautiful': 620,

'gets': 621,

'mean': 622,

'priced': 623,

'weekend': 624,

'easy': 625,

'stayed': 626,

'nthey': 627,

'life': 628,

'saying': 629,

'chinese': 630,

'nif': 631,

'months': 632,

'dirty': 633,

'sorry': 634,

'knew': 635,

"weren't": 636,

'needs': 637,

'remember': 638,

'cost': 639,

'salsa': 640,

'nso': 641,

'totally': 642,

'walking': 643,

'office': 644,

'main': 645,

'across': 646,

'12': 647,

'appetizer': 648,

'egg': 649,

'french': 650,

'seating': 651,

'potatoes': 652,

'view': 653,

'choice': 654,

'eggs': 655,

'phoenix': 656,

'near': 657,

'stopped': 658,

'available': 659,

'de': 660,

'please': 661,

'itself': 662,

'pasta': 663,

'orders': 664,

'hand': 665,

'bite': 666,

'type': 667,

'often': 668,

'bartender': 669,

'9': 670,

'read': 671,

'add': 672,

'attentive': 673,

'others': 674,

'la': 675,

'sign': 676,

'ordering': 677,

'italian': 678,

'paying': 679,

'patio': 680,

'flavors': 681,

'plenty': 682,

'beans': 683,

'potato': 684,

'due': 685,

'expected': 686,

'spend': 687,

'noticed': 688,

'hit': 689,

'sandwiches': 690,

'evening': 691,

'awful': 692,

'sunday': 693,

'taken': 694,

'says': 695,

'bowl': 696,

'watch': 697,

'reasonable': 698,

'25': 699,

'chef': 700,

"you'll": 701,

'servers': 702,

'head': 703,

'garlic': 704,

'perfectly': 705,

'yummy': 706,

'unfortunately': 707,

'along': 708,

'portion': 709,

'literally': 710,

'giving': 711,

'number': 712,

'simple': 713,

'black': 714,

'talk': 715,

'none': 716,

'idea': 717,

'five': 718,

'wow': 719,

'girls': 720,

'friday': 721,

'weeks': 722,

'variety': 723,

'eaten': 724,

'loud': 725,

'checked': 726,

'overpriced': 727,

'appointment': 728,

'birthday': 729,

'unless': 730,

'mediocre': 731,

'grilled': 732,

'together': 733,

'minute': 734,

'bought': 735,

'oil': 736,

'charged': 737,

'game': 738,

'bathroom': 739,

'boyfriend': 740,

'closed': 741,

'hope': 742,

'opened': 743,

'packed': 744,

'spent': 745,

'reservation': 746,

'yourself': 747,

'total': 748,

'credit': 749,

'talking': 750,

'taco': 751,

'turned': 752,

'month': 753,

'excited': 754,

'happened': 755,

'low': 756,

'hungry': 757,

'piece': 758,

'longer': 759,

'space': 760,

'surprised': 761,

'management': 762,

'immediately': 763,

'moved': 764,

'thank': 765,

'40': 766,

'pieces': 767,

'butter': 768,

'nit': 769,

'city': 770,

'comfortable': 771,

'hostess': 772,

'seafood': 773,

"haven't": 774,

'bottle': 775,

'issue': 776,

'ones': 777,

'lobster': 778,

'45': 779,

'crispy': 780,

'based': 781,

'per': 782,

'located': 783,

'choose': 784,

'seat': 785,

'sometimes': 786,

'salmon': 787,

'middle': 788,

'hear': 789,

'cup': 790,

'bed': 791,

'thanks': 792,

'medium': 793,

'waste': 794,

'within': 795,

'plates': 796,

'disappointing': 797,

'crowd': 798,

'somewhere': 799,

'prepared': 800,

'frozen': 801,

'case': 802,

'simply': 803,

'world': 804,

'noodles': 805,

'feeling': 806,

'glad': 807,

'worse': 808,

'beers': 809,

'hands': 810,

'chance': 811,

'including': 812,

'attitude': 813,

'play': 814,

'ingredients': 815,

'recently': 816,

'goes': 817,

'mention': 818,

'asking': 819,

'supposed': 820,

'mouth': 821,

'park': 822,

'center': 823,

'burrito': 824,

'fix': 825,

'onion': 826,

'salty': 827,

'exactly': 828,

'cafe': 829,

'upon': 830,

'w': 831,

'basically': 832,

'filled': 833,

'11': 834,

'stand': 835,

'meals': 836,

"aren't": 837,

'crowded': 838,

'anywhere': 839,

'brunch': 840,

'takes': 841,

'shopping': 842,

'showed': 843,

'recommended': 844,

'except': 845,

'pho': 846,

'nwhen': 847,

'move': 848,

'noverall': 849,

'hate': 850,

'staying': 851,

'99': 852,

'honestly': 853,

'scottsdale': 854,

'sad': 855,

'attention': 856,

'choices': 857,

'whatever': 858,

'100': 859,

'wall': 860,

'lost': 861,

'added': 862,

'box': 863,

'ambiance': 864,

'option': 865,

'appetizers': 866,

'mac': 867,

'apparently': 868,

'changed': 869,

'date': 870,

'lack': 871,

'tiny': 872,

'thinking': 873,

'pricey': 874,

'anyway': 875,

'avoid': 876,

'write': 877,

'save': 878,

'mall': 879,

'sent': 880,

'crust': 881,

'mine': 882,

'interesting': 883,

'finish': 884,

'tomato': 885,

'specials': 886,

'worked': 887,

'face': 888,

'pleasant': 889,

'flavorful': 890,

'asian': 891,

'weird': 892,

'authentic': 893,

'online': 894,

'cute': 895,

'cook': 896,

'dark': 897,

'cannot': 898,

'tuna': 899,

'delivery': 900,

'professional': 901,

'cash': 902,

'single': 903,

'soft': 904,

'level': 905,

'salt': 906,

'woman': 907,

'forward': 908,

'above': 909,

'rib': 910,

'barely': 911,

'finished': 912,

'sausage': 913,

'eye': 914,

'greeted': 915,

'dollars': 916,

'thin': 917,

'greasy': 918,

'tender': 919,

'bucks': 920,

'obviously': 921,

'sides': 922,

'bag': 923,

'sort': 924,

'note': 925,

'nails': 926,

'downtown': 927,

'tasting': 928,

'standard': 929,

'item': 930,

'min': 931,

'massage': 932,

'dance': 933,

'lettuce': 934,

'clearly': 935,

'seats': 936,

'running': 937,

'serving': 938,

'dr': 939,

'tv': 940,

'leaving': 941,

'u00e9': 942,

'miss': 943,

'crazy': 944,

'corn': 945,

'ribs': 946,

'forgot': 947,

'dogs': 948,

'onions': 949,

'unique': 950,

'school': 951,

'picked': 952,

'truly': 953,

'matter': 954,

'airport': 955,

'treat': 956,

'yeah': 957,

'turn': 958,

'mix': 959,

'joint': 960,

'dressing': 961,

'curry': 962,

'share': 963,

'book': 964,

'entrees': 965,

'establishment': 966,

'nservice': 967,

'dont': 968,

'section': 969,

'young': 970,

'groupon': 971,

'rare': 972,

'salads': 973,

'elsewhere': 974,

'pass': 975,

'speak': 976,

'fair': 977,

'expecting': 978,

'valley': 979,

'salon': 980,

"we're": 981,

'gross': 982,

'employee': 983,

'certainly': 984,

'b': 985,

'playing': 986,

'value': 987,

'tastes': 988,

'ridiculous': 989,

'nas': 990,

'nthere': 991,

'mentioned': 992,

'fee': 993,

'station': 994,

'veggies': 995,

'afternoon': 996,

'entree': 997,

'window': 998,

'using': 999,

'issues': 1000,

...}

sequences = tokenizer.texts\_to\_sequences(reviews)

index = 42

print(reviews[index])

print(sequences[index])

Some of the worst pizza I've ever had. We used a coupon from the paper for a 2 topping 8 cut Sicilian. First of all the pizza wasn't even cut through, and the sad attempt at cutting was so uneven that 4 of the slices were about an inch wide, while the others were about 4\" each. The toppings were scarce, they used mini pepperoni and put maybe 8 on the whole pizza. The onions were huge chunks and the mushrooms were straight from a can. The worst part though was the thick doughy crust that tasted more like a fishy sourdough roll. I'm serious... It was so noticeable that it made me wonder if the dough was bad or if they for some weird reason put fish sauce in it. It was gross. \n\nWe also ordered steak and Italian hoagies. The veggies were old and wilted, and there was no dressing on either. The Italian had deli meat that was clearly bottom of the line and not very generous. The \"steak\" (if you an call it that) was greyish instead of brown and looked like it was a processed meat chopped into pieces. No flavor or seasoning and the texture was reminiscent of spam. It was so bad that I only ate 1/4 of it and tossed the rest. \n\nI have ordered from here in the past and always been disappointed. I thought I would give them another try since I'd never ordered a Sicilian pizza from there. What a mistake. I will never order from them again!

[67, 7, 1, 384, 171, 112, 151, 24, 15, 299, 5, 1073, 50, 1, 1090, 9, 5, 115, 2403, 479, 501, 7487, 105, 7, 37, 1, 171, 168, 74, 501, 292, 2, 1, 855, 1992, 25, 2543, 6, 27, 4979, 13, 202, 7, 1, 1247, 26, 56, 59, 2584, 1712, 166, 1, 674, 26, 56, 202, 304, 1, 1027, 26, 9651, 19, 299, 1301, 2150, 2, 271, 273, 479, 21, 1, 375, 171, 1, 949, 26, 323, 2138, 2, 1, 1174, 26, 1193, 50, 5, 75, 1, 384, 377, 195, 6, 1, 1076, 4608, 881, 13, 321, 73, 38, 5, 3163, 4442, 527, 93, 1610, 8, 6, 27, 6270, 13, 8, 138, 32, 1459, 39, 1, 2046, 6, 143, 49, 39, 19, 9, 67, 892, 435, 271, 383, 181, 10, 8, 8, 6, 982, 12, 261, 84, 89, 309, 2, 678, 1, 995, 26, 254, 2, 4716, 2, 33, 6, 57, 961, 21, 405, 1, 678, 24, 1726, 258, 13, 6, 935, 1150, 7, 1, 260, 2, 22, 42, 1349, 1, 309, 39, 20, 59, 290, 8, 13, 6, 389, 7, 1285, 2, 269, 38, 8, 6, 5, 3871, 258, 2196, 178, 767, 57, 279, 49, 1789, 2, 1, 1280, 6, 5415, 7, 5604, 8, 6, 27, 143, 13, 3, 68, 506, 200, 202, 7, 8, 2, 3000, 1, 615, 12, 85, 23, 89, 50, 45, 10, 1, 546, 2, 128, 69, 337, 3, 252, 3, 53, 163, 76, 157, 121, 167, 364, 88, 89, 5, 7487, 171, 50, 33, 60, 5, 1002, 3, 66, 88, 96, 50, 76, 108]

print(tokenizer.word\_index['some'])

print(tokenizer.word\_index['of'])

print(tokenizer.word\_index['the'])

print(tokenizer.word\_index['worst'])

67

7

1

384

x\_train = pad\_sequences(sequences, maxlen=max\_review\_len)

x\_train[0]

array([ 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 707, 1, 4339, 7, 173, 939, 1665, 11, 5,

2161, 7, 1, 137, 112, 24, 18, 27, 189, 80, 3128,

10, 2636, 34, 1416, 462, 129, 8, 519, 13, 177, 129,

803, 88, 3812, 1, 446, 8, 392, 841, 115, 360, 7,

4136, 1438, 4, 43, 59, 1266, 118, 95, 46, 9, 13,

49, 1687, 4, 418, 18, 8, 3, 23, 563, 178, 16,

460, 18, 189, 80, 3128, 2, 3, 40, 78, 43, 8,

20, 23, 644, 1343, 20, 23, 3252, 18, 2921, 637, 227,

430, 476, 4312, 1, 446, 63, 2, 22, 215, 1, 63,

18, 2269, 13, 3, 249, 13, 3, 23, 4, 163, 939,

115, 237], dtype=int32)

model = Sequential()

model.add(Embedding(num\_words, 64, input\_length=max\_review\_len))

model.add(GRU(128))

model.add(Dense(1, activation='sigmoid'))

model.compile(optimizer='adam',

loss='binary\_crossentropy',

metrics=['accuracy'])

model\_save\_path = '/content/best\_model.h5'

checkpoint\_callback = ModelCheckpoint(model\_save\_path,

monitor='val\_accuracy',

save\_best\_only=True,

verbose=1)

history = model.fit(x\_train,

y\_train,

epochs=5,

batch\_size=128,

validation\_split=0.1,

callbacks=[checkpoint\_callback])

Epoch 1/5

3938/3938 [==============================] - ETA: 0s - loss: 0.1882 - accuracy: 0.9245

Epoch 00001: val\_accuracy improved from -inf to 0.95968, saving model to /content/best\_model.h5

3938/3938 [==============================] - 164s 40ms/step - loss: 0.1882 - accuracy: 0.9245 - val\_loss: 0.1060 - val\_accuracy: 0.9597

Epoch 2/5

3937/3938 [============================>.] - ETA: 0s - loss: 0.1185 - accuracy: 0.9546

Epoch 00002: val\_accuracy improved from 0.95968 to 0.96548, saving model to /content/best\_model.h5

3938/3938 [==============================] - 158s 40ms/step - loss: 0.1185 - accuracy: 0.9546 - val\_loss: 0.0918 - val\_accuracy: 0.9655

Epoch 3/5

3937/3938 [============================>.] - ETA: 0s - loss: 0.0964 - accuracy: 0.9638

Epoch 00003: val\_accuracy did not improve from 0.96548

3938/3938 [==============================] - 158s 40ms/step - loss: 0.0964 - accuracy: 0.9638 - val\_loss: 0.0912 - val\_accuracy: 0.9654

Epoch 4/5

3937/3938 [============================>.] - ETA: 0s - loss: 0.0791 - accuracy: 0.9709

Epoch 00004: val\_accuracy did not improve from 0.96548

3938/3938 [==============================] - 158s 40ms/step - loss: 0.0792 - accuracy: 0.9709 - val\_loss: 0.0970 - val\_accuracy: 0.9643

Epoch 5/5

3938/3938 [==============================] - ETA: 0s - loss: 0.0633 - accuracy: 0.9772

Epoch 00005: val\_accuracy did not improve from 0.96548

3938/3938 [==============================] - 158s 40ms/step - loss: 0.0633 - accuracy: 0.9772 - val\_loss: 0.1061 - val\_accuracy: 0.9641

model.load\_weights(model\_save\_path)

test\_sequences = tokenizer.texts\_to\_sequences(test['Review'])

x\_test = pad\_sequences(test\_sequences, maxlen=max\_review\_len)

x\_test[0]

array([ 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 8158, 4, 80, 336, 3, 23, 1361, 1547, 56,

1, 52, 49, 1, 256, 3, 23, 69, 287, 1837, 52,

45, 9, 1, 546, 148, 274, 165, 2, 1221, 4, 14,

137, 18, 289, 38, 8346, 2311, 257, 511, 28, 1167, 2,

113, 60, 568, 520, 1426, 16, 11, 44, 31, 13, 3,

82, 22, 249, 38, 3, 144, 173, 694, 2235, 7, 40,

72, 7, 14, 80, 2540, 5951, 23, 69, 9, 21, 14,

7, 1558, 2, 23, 1696, 14, 1894, 1775, 470, 17, 45,

14, 52, 2, 1346, 5923, 95, 37, 69, 91, 1024, 2,

311, 48, 4, 32, 4, 1480, 1037, 19, 40, 4124, 1,

317, 127, 8, 562, 5, 203, 114, 99, 8, 90, 10,

1168, 274], dtype=int32)

scores = model.evaluate(x\_test, y\_test, verbose=1)

print("The percent of correct answers:", round(scores[1] \* 100, 4))

1188/1188 [==============================] - 20s 17ms/step - loss: 0.1138 - accuracy: 0.9569

The percent of correct answers: 95.6868

OUTPUT:

{{

1188/1188 [==============================] - 20s 17ms/step - loss: 0.1138 - accuracy: 0.9569

The percent of correct answers: 95.6868

}}