model\_hype = Namespace(

EMBEDDING\_SIZE = 64,

GRU\_UNITS = 128,

ATTN\_SIZE = 2,

)

Epoch 0 Batch 0 Loss1.3382

Epoch 0 Batch 100 Loss0.6427

Epoch 0 Batch 200 Loss0.5850

Epoch 0 Batch 300 Loss0.5772

Epoch 0 Batch 400 Loss0.5579

Epoch 0 Batch 500 Loss0.5694

Epoch 0 Batch 600 Loss0.4850

Epoch 1 Loss 0.6048

Time taken for 1 epoch 107.60845971107483 sec

Epoch 1 Batch 0 Loss0.5220

Epoch 1 Batch 100 Loss0.5383

Epoch 1 Batch 200 Loss0.5129

Epoch 1 Batch 300 Loss0.4784

Epoch 1 Batch 400 Loss0.4806

Epoch 1 Batch 500 Loss0.4401

Epoch 1 Batch 600 Loss0.4116

Epoch 2 Loss 0.4631

Time taken for 1 epoch 59.3316011428833 sec

Epoch 2 Batch 0 Loss0.3942

Epoch 2 Batch 100 Loss0.3962

Epoch 2 Batch 200 Loss0.4096

Epoch 2 Batch 300 Loss0.3629

Epoch 2 Batch 400 Loss0.4016

Epoch 2 Batch 500 Loss0.4223

Epoch 2 Batch 600 Loss0.3991

Epoch 3 Loss 0.3931

Time taken for 1 epoch 59.42119312286377 sec

Epoch 3 Batch 0 Loss0.3444

Epoch 3 Batch 100 Loss0.3521

Epoch 3 Batch 200 Loss0.3265

Epoch 3 Batch 300 Loss0.3441

Epoch 3 Batch 400 Loss0.3561

Epoch 3 Batch 500 Loss0.3673

Epoch 3 Batch 600 Loss0.3388

Epoch 4 Loss 0.3390

Time taken for 1 epoch 58.85515069961548 sec

Epoch 4 Batch 0 Loss0.2783

Epoch 4 Batch 100 Loss0.3501

Epoch 4 Batch 200 Loss0.3144

Epoch 4 Batch 300 Loss0.3076

Epoch 4 Batch 400 Loss0.2856

Epoch 4 Batch 500 Loss0.2809

Epoch 4 Batch 600 Loss0.3017

Epoch 5 Loss 0.2941

Time taken for 1 epoch 58.792253494262695 sec

Epoch 5 Batch 0 Loss0.2300

Epoch 5 Batch 100 Loss0.2547

Epoch 5 Batch 200 Loss0.3260

Epoch 5 Batch 300 Loss0.2446

Epoch 5 Batch 400 Loss0.2716

Epoch 5 Batch 500 Loss0.2921

Epoch 5 Batch 600 Loss0.2784

Epoch 6 Loss 0.2585

Time taken for 1 epoch 57.96512818336487 sec

Epoch 6 Batch 0 Loss0.1947

Epoch 6 Batch 100 Loss0.2356

Epoch 6 Batch 200 Loss0.2424

Epoch 6 Batch 300 Loss0.2389

Epoch 6 Batch 400 Loss0.2333

Epoch 6 Batch 500 Loss0.2050

Epoch 6 Batch 600 Loss0.2055

Epoch 7 Loss 0.2293

Time taken for 1 epoch 57.7753963470459 sec

Epoch 7 Batch 0 Loss0.1913

Epoch 7 Batch 100 Loss0.1992

Epoch 7 Batch 200 Loss0.2095

Epoch 7 Batch 300 Loss0.2176

Epoch 7 Batch 400 Loss0.2112

Epoch 7 Batch 500 Loss0.1875

Epoch 7 Batch 600 Loss0.1832

Epoch 8 Loss 0.2062

Time taken for 1 epoch 57.736313343048096 sec

Epoch 8 Batch 0 Loss0.1659

Epoch 8 Batch 100 Loss0.1398

Epoch 8 Batch 200 Loss0.2938

Epoch 8 Batch 300 Loss0.2173

Epoch 8 Batch 400 Loss0.2401

Epoch 8 Batch 500 Loss0.1963

Epoch 8 Batch 600 Loss0.2162

Epoch 9 Loss 0.1999

Time taken for 1 epoch 57.892983198165894 sec

Epoch 9 Batch 0 Loss0.1651

Epoch 9 Batch 100 Loss0.1345

Epoch 9 Batch 200 Loss0.1846

Epoch 9 Batch 300 Loss0.1820

Epoch 9 Batch 400 Loss0.1725

Epoch 9 Batch 500 Loss0.1734

Epoch 9 Batch 600 Loss0.1656

Epoch 10 Loss 0.1781

Time taken for 1 epoch 58.065853118896484 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 64,

    GRU\_UNITS = 512,

    ATTN\_SIZE = 2,

)

Epoch 0 Batch 0 Loss1.4004

Epoch 0 Batch 100 Loss0.5468

Epoch 0 Batch 200 Loss0.5463

Epoch 0 Batch 300 Loss0.4872

Epoch 0 Batch 400 Loss0.5090

Epoch 0 Batch 500 Loss0.3886

Epoch 0 Batch 600 Loss0.5827

Epoch 1 Loss 0.5514

Time taken for 1 epoch 168.24856281280518 sec

Epoch 1 Batch 0 Loss0.5546

Epoch 1 Batch 100 Loss0.5250

Epoch 1 Batch 200 Loss0.4952

Epoch 1 Batch 300 Loss0.4908

Epoch 1 Batch 400 Loss0.5098

Epoch 1 Batch 500 Loss0.5212

Epoch 1 Batch 600 Loss0.4817

Epoch 2 Loss 0.5029

Time taken for 1 epoch 119.81172347068787 sec

Epoch 2 Batch 0 Loss0.4157

Epoch 2 Batch 100 Loss0.4302

Epoch 2 Batch 200 Loss0.4014

Epoch 2 Batch 300 Loss0.4000

Epoch 2 Batch 400 Loss0.4085

Epoch 2 Batch 500 Loss0.4243

Epoch 2 Batch 600 Loss0.4614

Epoch 3 Loss 0.4312

Time taken for 1 epoch 119.85378575325012 sec

Epoch 3 Batch 0 Loss0.3906

Epoch 3 Batch 100 Loss0.3449

Epoch 3 Batch 200 Loss0.4063

Epoch 3 Batch 300 Loss0.4083

Epoch 3 Batch 400 Loss0.3942

Epoch 3 Batch 500 Loss0.4097

Epoch 3 Batch 600 Loss0.4046

Epoch 4 Loss 0.3949

Time taken for 1 epoch 119.7948899269104 sec

Epoch 4 Batch 0 Loss0.3755

Epoch 4 Batch 100 Loss0.3702

Epoch 4 Batch 200 Loss0.3392

Epoch 4 Batch 300 Loss0.3464

Epoch 4 Batch 400 Loss0.3250

Epoch 4 Batch 500 Loss0.4099

Epoch 4 Batch 600 Loss0.3960

Epoch 5 Loss 0.3673

Time taken for 1 epoch 119.51410579681396 sec

Epoch 5 Batch 0 Loss0.3633

Epoch 5 Batch 100 Loss0.3009

Epoch 5 Batch 200 Loss0.3244

Epoch 5 Batch 300 Loss0.3693

Epoch 5 Batch 400 Loss0.3309

Epoch 5 Batch 500 Loss0.3408

Epoch 5 Batch 600 Loss0.3960

Epoch 6 Loss 0.3465

Time taken for 1 epoch 119.61765718460083 sec

Epoch 6 Batch 0 Loss0.3142

Epoch 6 Batch 100 Loss0.3197

Epoch 6 Batch 200 Loss0.3263

Epoch 6 Batch 300 Loss0.3156

Epoch 6 Batch 400 Loss0.3548

Epoch 6 Batch 500 Loss0.3804

Epoch 6 Batch 600 Loss0.3233

Epoch 7 Loss 0.3315

Time taken for 1 epoch 119.38740086555481 sec

Epoch 7 Batch 0 Loss0.3260

Epoch 7 Batch 100 Loss0.2770

Epoch 7 Batch 200 Loss0.3311

Epoch 7 Batch 300 Loss0.3098

Epoch 7 Batch 400 Loss0.3374

Epoch 7 Batch 500 Loss0.3094

Epoch 7 Batch 600 Loss0.3036

Epoch 8 Loss 0.3181

Time taken for 1 epoch 119.54596543312073 sec

Epoch 8 Batch 0 Loss0.2747

Epoch 8 Batch 100 Loss0.3015

Epoch 8 Batch 200 Loss0.3483

Epoch 8 Batch 300 Loss0.2915

Epoch 8 Batch 400 Loss0.3202

Epoch 8 Batch 500 Loss0.3550

Epoch 8 Batch 600 Loss0.3560

Epoch 9 Loss 0.3050

Time taken for 1 epoch 119.34983086585999 sec

Epoch 9 Batch 0 Loss0.2837

Epoch 9 Batch 100 Loss0.3025

Epoch 9 Batch 200 Loss0.3025

Epoch 9 Batch 300 Loss0.2554

Epoch 9 Batch 400 Loss0.2981

Epoch 9 Batch 500 Loss0.2901

Epoch 9 Batch 600 Loss0.2380

Epoch 10 Loss 0.2915

Time taken for 1 epoch 119.50812649726868 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 32,

    GRU\_UNITS = 256,

    ATTN\_SIZE = 2,

)

Epoch 0 Batch 0 Loss1.4518

Epoch 0 Batch 100 Loss0.6873

Epoch 0 Batch 200 Loss0.5967

Epoch 0 Batch 300 Loss0.5486

Epoch 0 Batch 400 Loss0.5686

Epoch 0 Batch 500 Loss0.4945

Epoch 0 Batch 600 Loss0.5410

Epoch 1 Loss 0.6053

Time taken for 1 epoch 108.29345107078552 sec

Epoch 1 Batch 0 Loss0.5195

Epoch 1 Batch 100 Loss0.4826

Epoch 1 Batch 200 Loss0.4606

Epoch 1 Batch 300 Loss0.4800

Epoch 1 Batch 400 Loss0.4519

Epoch 1 Batch 500 Loss0.4012

Epoch 1 Batch 600 Loss0.4112

Epoch 2 Loss 0.4520

Time taken for 1 epoch 60.89512896537781 sec

Epoch 2 Batch 0 Loss0.4234

Epoch 2 Batch 100 Loss0.3844

Epoch 2 Batch 200 Loss0.3604

Epoch 2 Batch 300 Loss0.4218

Epoch 2 Batch 400 Loss0.3730

Epoch 2 Batch 500 Loss0.4228

Epoch 2 Batch 600 Loss0.4107

Epoch 3 Loss 0.3763

Time taken for 1 epoch 60.89751434326172 sec

Epoch 3 Batch 0 Loss0.3041

Epoch 3 Batch 100 Loss0.2994

Epoch 3 Batch 200 Loss0.3436

Epoch 3 Batch 300 Loss0.3405

Epoch 3 Batch 400 Loss0.2541

Epoch 3 Batch 500 Loss0.2733

Epoch 3 Batch 600 Loss0.3058

Epoch 4 Loss 0.3176

Time taken for 1 epoch 61.04163312911987 sec

Epoch 4 Batch 0 Loss0.2392

Epoch 4 Batch 100 Loss0.2678

Epoch 4 Batch 200 Loss0.3181

Epoch 4 Batch 300 Loss0.2594

Epoch 4 Batch 400 Loss0.2629

Epoch 4 Batch 500 Loss0.2535

Epoch 4 Batch 600 Loss0.2458

Epoch 5 Loss 0.2693

Time taken for 1 epoch 61.3652663230896 sec

Epoch 5 Batch 0 Loss0.2408

Epoch 5 Batch 100 Loss0.2135

Epoch 5 Batch 200 Loss0.2149

Epoch 5 Batch 300 Loss0.2452

Epoch 5 Batch 400 Loss0.2432

Epoch 5 Batch 500 Loss0.2172

Epoch 5 Batch 600 Loss0.2027

Epoch 6 Loss 0.2334

Time taken for 1 epoch 61.75162410736084 sec

Epoch 6 Batch 0 Loss0.1972

Epoch 6 Batch 100 Loss0.2190

Epoch 6 Batch 200 Loss0.2163

Epoch 6 Batch 300 Loss0.1939

Epoch 6 Batch 400 Loss0.2056

Epoch 6 Batch 500 Loss0.1764

Epoch 6 Batch 600 Loss0.1984

Epoch 7 Loss 0.2046

Time taken for 1 epoch 61.75383281707764 sec

Epoch 7 Batch 0 Loss0.1986

Epoch 7 Batch 100 Loss0.1690

Epoch 7 Batch 200 Loss0.2020

Epoch 7 Batch 300 Loss0.1636

Epoch 7 Batch 400 Loss0.2239

Epoch 7 Batch 500 Loss0.2107

Epoch 7 Batch 600 Loss0.2158

Epoch 8 Loss 0.1844

Time taken for 1 epoch 61.786009073257446 sec

Epoch 8 Batch 0 Loss0.1295

Epoch 8 Batch 100 Loss0.1483

Epoch 8 Batch 200 Loss0.1436

Epoch 8 Batch 300 Loss0.1781

Epoch 8 Batch 400 Loss0.1642

Epoch 8 Batch 500 Loss0.1428

Epoch 8 Batch 600 Loss0.1744

Epoch 9 Loss 0.1696

Time taken for 1 epoch 62.03496980667114 sec

Epoch 9 Batch 0 Loss0.1616

Epoch 9 Batch 100 Loss0.1502

Epoch 9 Batch 200 Loss0.1787

Epoch 9 Batch 300 Loss0.1539

Epoch 9 Batch 400 Loss0.1999

Epoch 9 Batch 500 Loss0.1539

Epoch 9 Batch 600 Loss0.1659

Epoch 10 Loss 0.1646

Time taken for 1 epoch 62.365535736083984 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 32,

    GRU\_UNITS = 256,

    ATTN\_SIZE = 5,

)

Epoch 0 Batch 0 Loss1.3463

Epoch 0 Batch 100 Loss0.7161

Epoch 0 Batch 200 Loss0.6332

Epoch 0 Batch 300 Loss0.6066

Epoch 0 Batch 400 Loss0.5931

Epoch 0 Batch 500 Loss0.5605

Epoch 0 Batch 600 Loss0.4732

Epoch 1 Loss 0.6342

Time taken for 1 epoch 112.72214126586914 sec

Epoch 1 Batch 0 Loss0.5189

Epoch 1 Batch 100 Loss0.5342

Epoch 1 Batch 200 Loss0.5345

Epoch 1 Batch 300 Loss0.4739

Epoch 1 Batch 400 Loss0.5081

Epoch 1 Batch 500 Loss0.5138

Epoch 1 Batch 600 Loss0.4768

Epoch 2 Loss 0.4976

Time taken for 1 epoch 61.22661781311035 sec

Epoch 2 Batch 0 Loss0.4884

Epoch 2 Batch 100 Loss0.4252

Epoch 2 Batch 200 Loss0.4429

Epoch 2 Batch 300 Loss0.4884

Epoch 2 Batch 400 Loss0.4430

Epoch 2 Batch 500 Loss0.4619

Epoch 2 Batch 600 Loss0.4359

Epoch 3 Loss 0.4467

Time taken for 1 epoch 61.56577467918396 sec

Epoch 3 Batch 0 Loss0.4242

Epoch 3 Batch 100 Loss0.4716

Epoch 3 Batch 200 Loss0.3989

Epoch 3 Batch 300 Loss0.4198

Epoch 3 Batch 400 Loss0.4373

Epoch 3 Batch 500 Loss0.4431

Epoch 3 Batch 600 Loss0.3528

Epoch 4 Loss 0.4139

Time taken for 1 epoch 61.71033239364624 sec

Epoch 4 Batch 0 Loss0.3932

Epoch 4 Batch 100 Loss0.3931

Epoch 4 Batch 200 Loss0.4032

Epoch 4 Batch 300 Loss0.4029

Epoch 4 Batch 400 Loss0.3512

Epoch 4 Batch 500 Loss0.3690

Epoch 4 Batch 600 Loss0.3545

Epoch 5 Loss 0.3869

Time taken for 1 epoch 61.50132966041565 sec

Epoch 5 Batch 0 Loss0.3601

Epoch 5 Batch 100 Loss0.3348

Epoch 5 Batch 200 Loss0.3547

Epoch 5 Batch 300 Loss0.3405

Epoch 5 Batch 400 Loss0.3718

Epoch 5 Batch 500 Loss0.3820

Epoch 5 Batch 600 Loss0.3516

Epoch 6 Loss 0.3588

Time taken for 1 epoch 61.30576014518738 sec

Epoch 6 Batch 0 Loss0.3255

Epoch 6 Batch 100 Loss0.3210

Epoch 6 Batch 200 Loss0.3203

Epoch 6 Batch 300 Loss0.3331

Epoch 6 Batch 400 Loss0.3216

Epoch 6 Batch 500 Loss0.2663

Epoch 6 Batch 600 Loss0.3248

Epoch 7 Loss 0.3296

Time taken for 1 epoch 60.89201235771179 sec

Epoch 7 Batch 0 Loss0.2638

Epoch 7 Batch 100 Loss0.3041

Epoch 7 Batch 200 Loss0.3687

Epoch 7 Batch 300 Loss0.3020

Epoch 7 Batch 400 Loss0.2832

Epoch 7 Batch 500 Loss0.3133

Epoch 7 Batch 600 Loss0.3035

Epoch 8 Loss 0.3031

Time taken for 1 epoch 60.88247799873352 sec

Epoch 8 Batch 0 Loss0.2708

Epoch 8 Batch 100 Loss0.2849

Epoch 8 Batch 200 Loss0.2519

Epoch 8 Batch 300 Loss0.3143

Epoch 8 Batch 400 Loss0.3145

Epoch 8 Batch 500 Loss0.3145

Epoch 8 Batch 600 Loss0.2721

Epoch 9 Loss 0.2786

Time taken for 1 epoch 60.81625699996948 sec

Epoch 9 Batch 0 Loss0.2284

Epoch 9 Batch 100 Loss0.2598

Epoch 9 Batch 200 Loss0.2370

Epoch 9 Batch 300 Loss0.2678

Epoch 9 Batch 400 Loss0.2299

Epoch 9 Batch 500 Loss0.2526

Epoch 9 Batch 600 Loss0.2649

Epoch 10 Loss 0.2576

Time taken for 1 epoch 60.6395206451416 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 32,

    GRU\_UNITS = 128,

    ATTN\_SIZE = 2,

)

Epoch 0 Batch 0 Loss1.3355

Epoch 0 Batch 100 Loss0.6732

Epoch 0 Batch 200 Loss0.6797

Epoch 0 Batch 300 Loss0.6278

Epoch 0 Batch 400 Loss0.5528

Epoch 0 Batch 500 Loss0.5560

Epoch 0 Batch 600 Loss0.4959

Epoch 1 Loss 0.6105

Time taken for 1 epoch 104.18795609474182 sec

Epoch 1 Batch 0 Loss0.4565

Epoch 1 Batch 100 Loss0.4656

Epoch 1 Batch 200 Loss0.4690

Epoch 1 Batch 300 Loss0.3987

Epoch 1 Batch 400 Loss0.4291

Epoch 1 Batch 500 Loss0.4750

Epoch 1 Batch 600 Loss0.4100

Epoch 2 Loss 0.4457

Time taken for 1 epoch 56.11178660392761 sec

Epoch 2 Batch 0 Loss0.3592

Epoch 2 Batch 100 Loss0.4177

Epoch 2 Batch 200 Loss0.3907

Epoch 2 Batch 300 Loss0.4342

Epoch 2 Batch 400 Loss0.3300

Epoch 2 Batch 500 Loss0.3480

Epoch 2 Batch 600 Loss0.3738

Epoch 3 Loss 0.3791

Time taken for 1 epoch 55.751290798187256 sec

Epoch 3 Batch 0 Loss0.3492

Epoch 3 Batch 100 Loss0.3216

Epoch 3 Batch 200 Loss0.3272

Epoch 3 Batch 300 Loss0.3189

Epoch 3 Batch 400 Loss0.3191

Epoch 3 Batch 500 Loss0.3671

Epoch 3 Batch 600 Loss0.2887

Epoch 4 Loss 0.3270

Time taken for 1 epoch 55.99023675918579 sec

Epoch 4 Batch 0 Loss0.2834

Epoch 4 Batch 100 Loss0.2717

Epoch 4 Batch 200 Loss0.2950

Epoch 4 Batch 300 Loss0.3002

Epoch 4 Batch 400 Loss0.2718

Epoch 4 Batch 500 Loss0.3022

Epoch 4 Batch 600 Loss0.2449

Epoch 5 Loss 0.2852

Time taken for 1 epoch 55.91382694244385 sec

Epoch 5 Batch 0 Loss0.2388

Epoch 5 Batch 100 Loss0.2759

Epoch 5 Batch 200 Loss0.2628

Epoch 5 Batch 300 Loss0.1986

Epoch 5 Batch 400 Loss0.2652

Epoch 5 Batch 500 Loss0.2550

Epoch 5 Batch 600 Loss0.2592

Epoch 6 Loss 0.2515

Time taken for 1 epoch 55.696765661239624 sec

Epoch 6 Batch 0 Loss0.2161

Epoch 6 Batch 100 Loss0.2385

Epoch 6 Batch 200 Loss0.2049

Epoch 6 Batch 300 Loss0.2463

Epoch 6 Batch 400 Loss0.2061

Epoch 6 Batch 500 Loss0.2289

Epoch 6 Batch 600 Loss0.2261

Epoch 7 Loss 0.2265

Time taken for 1 epoch 55.095093727111816 sec

Epoch 7 Batch 0 Loss0.1820

Epoch 7 Batch 100 Loss0.2033

Epoch 7 Batch 200 Loss0.2010

Epoch 7 Batch 300 Loss0.1677

Epoch 7 Batch 400 Loss0.2324

Epoch 7 Batch 500 Loss0.2427

Epoch 7 Batch 600 Loss0.2032

Epoch 8 Loss 0.2011

Time taken for 1 epoch 55.155548334121704 sec

Epoch 8 Batch 0 Loss0.1698

Epoch 8 Batch 100 Loss0.1835

Epoch 8 Batch 200 Loss0.2030

Epoch 8 Batch 300 Loss0.1722

Epoch 8 Batch 400 Loss0.2289

Epoch 8 Batch 500 Loss0.1405

Epoch 8 Batch 600 Loss0.2010

Epoch 9 Loss 0.1789

Time taken for 1 epoch 55.07136368751526 sec

Epoch 9 Batch 0 Loss0.1389

Epoch 9 Batch 100 Loss0.1460

Epoch 9 Batch 200 Loss0.1703

Epoch 9 Batch 300 Loss0.1876

Epoch 9 Batch 400 Loss0.1631

Epoch 9 Batch 500 Loss0.1836

Epoch 9 Batch 600 Loss0.1534

Epoch 10 Loss 0.1661

Time taken for 1 epoch 54.68582487106323 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 16,

    GRU\_UNITS = 128,

    ATTN\_SIZE = 2,

)

Epoch 0 Batch 0 Loss1.2841

Epoch 0 Batch 100 Loss0.7474

Epoch 0 Batch 200 Loss0.6901

Epoch 0 Batch 300 Loss0.6168

Epoch 0 Batch 400 Loss0.5842

Epoch 0 Batch 500 Loss0.5310

Epoch 0 Batch 600 Loss0.5627

Epoch 1 Loss 0.6418

Time taken for 1 epoch 103.69138860702515 sec

Epoch 1 Batch 0 Loss0.5345

Epoch 1 Batch 100 Loss0.5561

Epoch 1 Batch 200 Loss0.4764

Epoch 1 Batch 300 Loss0.4758

Epoch 1 Batch 400 Loss0.4578

Epoch 1 Batch 500 Loss0.5139

Epoch 1 Batch 600 Loss0.4148

Epoch 2 Loss 0.4668

Time taken for 1 epoch 54.286051750183105 sec

Epoch 2 Batch 0 Loss0.4024

Epoch 2 Batch 100 Loss0.3772

Epoch 2 Batch 200 Loss0.4119

Epoch 2 Batch 300 Loss0.3796

Epoch 2 Batch 400 Loss0.3866

Epoch 2 Batch 500 Loss0.4030

Epoch 2 Batch 600 Loss0.3599

Epoch 3 Loss 0.3844

Time taken for 1 epoch 54.22299313545227 sec

Epoch 3 Batch 0 Loss0.3364

Epoch 3 Batch 100 Loss0.3208

Epoch 3 Batch 200 Loss0.3058

Epoch 3 Batch 300 Loss0.2883

Epoch 3 Batch 400 Loss0.2910

Epoch 3 Batch 500 Loss0.3174

Epoch 3 Batch 600 Loss0.3110

Epoch 4 Loss 0.3217

Time taken for 1 epoch 53.9471960067749 sec

Epoch 4 Batch 0 Loss0.3414

Epoch 4 Batch 100 Loss0.2699

Epoch 4 Batch 200 Loss0.2387

Epoch 4 Batch 300 Loss0.2851

Epoch 4 Batch 400 Loss0.2765

Epoch 4 Batch 500 Loss0.2540

Epoch 4 Batch 600 Loss0.2025

Epoch 5 Loss 0.2720

Time taken for 1 epoch 53.96268343925476 sec

Epoch 5 Batch 0 Loss0.2123

Epoch 5 Batch 100 Loss0.2392

Epoch 5 Batch 200 Loss0.2399

Epoch 5 Batch 300 Loss0.2109

Epoch 5 Batch 400 Loss0.2477

Epoch 5 Batch 500 Loss0.2365

Epoch 5 Batch 600 Loss0.2077

Epoch 6 Loss 0.2330

Time taken for 1 epoch 53.89530324935913 sec

Epoch 6 Batch 0 Loss0.2063

Epoch 6 Batch 100 Loss0.2067

Epoch 6 Batch 200 Loss0.2142

Epoch 6 Batch 300 Loss0.2013

Epoch 6 Batch 400 Loss0.1734

Epoch 6 Batch 500 Loss0.1882

Epoch 6 Batch 600 Loss0.2014

Epoch 7 Loss 0.2022

Time taken for 1 epoch 53.94658279418945 sec

Epoch 7 Batch 0 Loss0.1386

Epoch 7 Batch 100 Loss0.1835

Epoch 7 Batch 200 Loss0.1784

Epoch 7 Batch 300 Loss0.1989

Epoch 7 Batch 400 Loss0.1859

Epoch 7 Batch 500 Loss0.1450

Epoch 7 Batch 600 Loss0.1775

Epoch 8 Loss 0.1769

Time taken for 1 epoch 53.65388584136963 sec

Epoch 8 Batch 0 Loss0.1385

Epoch 8 Batch 100 Loss0.1359

Epoch 8 Batch 200 Loss0.1528

Epoch 8 Batch 300 Loss0.1533

Epoch 8 Batch 400 Loss0.1585

Epoch 8 Batch 500 Loss0.1489

Epoch 8 Batch 600 Loss0.2044

Epoch 9 Loss 0.1581

Time taken for 1 epoch 53.694849252700806 sec

Epoch 9 Batch 0 Loss0.1437

Epoch 9 Batch 100 Loss0.1539

Epoch 9 Batch 200 Loss0.1081

Epoch 9 Batch 300 Loss0.1166

Epoch 9 Batch 400 Loss0.1439

Epoch 9 Batch 500 Loss0.1392

Epoch 9 Batch 600 Loss0.1472

Epoch 10 Loss 0.1437

Time taken for 1 epoch 53.55032801628113 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 256,

    GRU\_UNITS = 1024,

    ATTN\_SIZE = 10,

)

Epoch 0 Batch 0 Loss1.4058

Epoch 0 Batch 100 Loss0.7203

Epoch 0 Batch 200 Loss0.6250

Epoch 0 Batch 300 Loss0.6136

Epoch 0 Batch 400 Loss0.6328

Epoch 0 Batch 500 Loss0.6548

Epoch 0 Batch 600 Loss0.6158

Epoch 1 Loss 0.7236

Time taken for 1 epoch 368.96156454086304 sec

Epoch 1 Batch 0 Loss0.6006

Epoch 1 Batch 100 Loss0.5938

Epoch 1 Batch 200 Loss0.5802

Epoch 1 Batch 300 Loss0.6372

Epoch 1 Batch 400 Loss0.6013

Epoch 1 Batch 500 Loss0.5850

Epoch 1 Batch 600 Loss0.5507

Epoch 2 Loss 0.5878

Time taken for 1 epoch 321.4963903427124 sec

Epoch 2 Batch 0 Loss0.5552

Epoch 2 Batch 100 Loss0.5839

Epoch 2 Batch 200 Loss0.5691

Epoch 2 Batch 300 Loss0.5782

Epoch 2 Batch 400 Loss0.6407

Epoch 2 Batch 500 Loss0.5284

Epoch 2 Batch 600 Loss0.5362

Epoch 3 Loss 0.5763

Time taken for 1 epoch 320.84988164901733 sec

Epoch 3 Batch 0 Loss0.6317

Epoch 3 Batch 100 Loss0.5439

Epoch 3 Batch 200 Loss0.5750

Epoch 3 Batch 300 Loss0.5705

Epoch 3 Batch 400 Loss0.5380

Epoch 3 Batch 500 Loss0.5058

Epoch 3 Batch 600 Loss0.5530

Epoch 4 Loss 0.5578

Time taken for 1 epoch 321.06647968292236 sec

Epoch 4 Batch 0 Loss0.5624

Epoch 4 Batch 100 Loss0.5358

Epoch 4 Batch 200 Loss0.4978

Epoch 4 Batch 300 Loss0.5252

Epoch 4 Batch 400 Loss0.5313

Epoch 4 Batch 500 Loss0.5019

Epoch 4 Batch 600 Loss0.5509

Epoch 5 Loss 0.5531

Time taken for 1 epoch 320.661673784256 sec

Epoch 5 Batch 0 Loss0.5124

Epoch 5 Batch 100 Loss0.6038

Epoch 5 Batch 200 Loss0.5123

Epoch 5 Batch 300 Loss0.5303

Epoch 5 Batch 400 Loss0.5220

Epoch 5 Batch 500 Loss0.5329

Epoch 5 Batch 600 Loss0.5237

Epoch 6 Loss 0.5445

Time taken for 1 epoch 320.76361060142517 sec

Epoch 6 Batch 0 Loss0.6078

Epoch 6 Batch 100 Loss0.5600

Epoch 6 Batch 200 Loss0.5001

Epoch 6 Batch 300 Loss0.5126

Epoch 6 Batch 400 Loss0.5177

Epoch 6 Batch 500 Loss0.4742

Epoch 6 Batch 600 Loss0.5260

Epoch 7 Loss 0.5334

Time taken for 1 epoch 320.12443017959595 sec

Epoch 7 Batch 0 Loss0.5364

Epoch 7 Batch 100 Loss0.4983

Epoch 7 Batch 200 Loss0.4804

Epoch 7 Batch 300 Loss0.4870

Epoch 7 Batch 400 Loss0.5752

Epoch 7 Batch 500 Loss0.5639

Epoch 7 Batch 600 Loss0.5538

Epoch 8 Loss 0.5233

Time taken for 1 epoch 320.7858974933624 sec

Epoch 8 Batch 0 Loss0.5235

Epoch 8 Batch 100 Loss0.5799

Epoch 8 Batch 200 Loss0.5186

Epoch 8 Batch 300 Loss0.5468

Epoch 8 Batch 400 Loss0.5162

Epoch 8 Batch 500 Loss0.5321

Epoch 8 Batch 600 Loss0.5255

Epoch 9 Loss 0.5250

Time taken for 1 epoch 320.22981882095337 sec

Epoch 9 Batch 0 Loss0.5152

Epoch 9 Batch 100 Loss0.5247

Epoch 9 Batch 200 Loss0.5311

Epoch 9 Batch 300 Loss0.5257

Epoch 9 Batch 400 Loss0.4693

Epoch 9 Batch 500 Loss0.5283

Epoch 9 Batch 600 Loss0.5411

Epoch 10 Loss 0.5313

Time taken for 1 epoch 320.6415021419525 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 64,

    GRU\_UNITS = 256,

    ATTN\_SIZE = 2,

)

Epoch 0 Batch 0 Loss1.3788

Epoch 0 Batch 100 Loss0.6749

Epoch 0 Batch 200 Loss0.6860

Epoch 0 Batch 300 Loss0.5059

Epoch 0 Batch 400 Loss0.5740

Epoch 0 Batch 500 Loss0.4795

Epoch 0 Batch 600 Loss0.4783

Epoch 1 Loss 0.5956

Time taken for 1 epoch 114.56185746192932 sec

Epoch 1 Batch 0 Loss0.3876

Epoch 1 Batch 100 Loss0.4744

Epoch 1 Batch 200 Loss0.4097

Epoch 1 Batch 300 Loss0.3784

Epoch 1 Batch 400 Loss0.4615

Epoch 1 Batch 500 Loss0.3785

Epoch 1 Batch 600 Loss0.3514

Epoch 2 Loss 0.4331

Time taken for 1 epoch 65.5784273147583 sec

Epoch 2 Batch 0 Loss0.3445

Epoch 2 Batch 100 Loss0.3727

Epoch 2 Batch 200 Loss0.3245

Epoch 2 Batch 300 Loss0.3574

Epoch 2 Batch 400 Loss0.3582

Epoch 2 Batch 500 Loss0.3862

Epoch 2 Batch 600 Loss0.4621

Epoch 3 Loss 0.3547

Time taken for 1 epoch 66.66680359840393 sec

Epoch 3 Batch 0 Loss0.3200

Epoch 3 Batch 100 Loss0.3004

Epoch 3 Batch 200 Loss0.3288

Epoch 3 Batch 300 Loss0.3296

Epoch 3 Batch 400 Loss0.2983

Epoch 3 Batch 500 Loss0.2996

Epoch 3 Batch 600 Loss0.3067

Epoch 4 Loss 0.2984

Time taken for 1 epoch 67.41628074645996 sec

Epoch 4 Batch 0 Loss0.3211

Epoch 4 Batch 100 Loss0.2649

Epoch 4 Batch 200 Loss0.2883

Epoch 4 Batch 300 Loss0.3069

Epoch 4 Batch 400 Loss0.2613

Epoch 4 Batch 500 Loss0.2974

Epoch 4 Batch 600 Loss0.1972

Epoch 5 Loss 0.2556

Time taken for 1 epoch 67.33098030090332 sec

Epoch 5 Batch 0 Loss0.1878

Epoch 5 Batch 100 Loss0.2111

Epoch 5 Batch 200 Loss0.2182

Epoch 5 Batch 300 Loss0.2882

Epoch 5 Batch 400 Loss0.2363

Epoch 5 Batch 500 Loss0.2491

Epoch 5 Batch 600 Loss0.1901

Epoch 6 Loss 0.2219

Time taken for 1 epoch 67.18223261833191 sec

Epoch 6 Batch 0 Loss0.2132

Epoch 6 Batch 100 Loss0.1956

Epoch 6 Batch 200 Loss0.2139

Epoch 6 Batch 300 Loss0.2577

Epoch 6 Batch 400 Loss0.2110

Epoch 6 Batch 500 Loss0.2208

Epoch 6 Batch 600 Loss0.2107

Epoch 7 Loss 0.2012

Time taken for 1 epoch 67.1350519657135 sec

Epoch 7 Batch 0 Loss0.1516

Epoch 7 Batch 100 Loss0.1369

Epoch 7 Batch 200 Loss0.1819

Epoch 7 Batch 300 Loss0.1562

Epoch 7 Batch 400 Loss0.2307

Epoch 7 Batch 500 Loss0.1471

Epoch 7 Batch 600 Loss0.1824

Epoch 8 Loss 0.1826

Time taken for 1 epoch 65.24883437156677 sec

Epoch 8 Batch 0 Loss0.1802

Epoch 8 Batch 100 Loss0.1704

Epoch 8 Batch 200 Loss0.1701

Epoch 8 Batch 300 Loss0.1501

Epoch 8 Batch 400 Loss0.1106

Epoch 8 Batch 500 Loss0.1959

Epoch 8 Batch 600 Loss0.1708

Epoch 9 Loss 0.1712

Time taken for 1 epoch 65.02187848091125 sec

Epoch 9 Batch 0 Loss0.1412

Epoch 9 Batch 100 Loss0.1604

Epoch 9 Batch 200 Loss0.1528

Epoch 9 Batch 300 Loss0.1454

Epoch 9 Batch 400 Loss0.1446

Epoch 9 Batch 500 Loss0.1622

Epoch 9 Batch 600 Loss0.2007

Epoch 10 Loss 0.1566

Time taken for 1 epoch 65.81846451759338 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 64,

    GRU\_UNITS = 128,

    ATTN\_SIZE = 2,

)



Epoch 0 Batch 0 Loss1.4112

Epoch 0 Batch 100 Loss0.6440

Epoch 0 Batch 200 Loss0.6811

Epoch 0 Batch 300 Loss0.5992

Epoch 0 Batch 400 Loss0.5884

Epoch 0 Batch 500 Loss0.6297

Epoch 0 Batch 600 Loss0.5870

Epoch 1 Loss 0.6194

Time taken for 1 epoch 112.74817276000977 sec

Epoch 1 Batch 0 Loss0.5787

Epoch 1 Batch 100 Loss0.5459

Epoch 1 Batch 200 Loss0.4954

Epoch 1 Batch 300 Loss0.4991

Epoch 1 Batch 400 Loss0.5455

Epoch 1 Batch 500 Loss0.5168

Epoch 1 Batch 600 Loss0.4818

Epoch 2 Loss 0.5267

Time taken for 1 epoch 63.306623220443726 sec

Epoch 2 Batch 0 Loss0.4411

Epoch 2 Batch 100 Loss0.4773

Epoch 2 Batch 200 Loss0.5016

Epoch 2 Batch 300 Loss0.4882

Epoch 2 Batch 400 Loss0.5399

Epoch 2 Batch 500 Loss0.5052

Epoch 2 Batch 600 Loss0.4901

Epoch 3 Loss 0.4817

Time taken for 1 epoch 61.80428338050842 sec

Epoch 3 Batch 0 Loss0.4224

Epoch 3 Batch 100 Loss0.4301

Epoch 3 Batch 200 Loss0.4746

Epoch 3 Batch 300 Loss0.4431

Epoch 3 Batch 400 Loss0.4955

Epoch 3 Batch 500 Loss0.4960

Epoch 3 Batch 600 Loss0.4340

Epoch 4 Loss 0.4541

Time taken for 1 epoch 62.73956561088562 sec

Epoch 4 Batch 0 Loss0.4381

Epoch 4 Batch 100 Loss0.4256

Epoch 4 Batch 200 Loss0.3963

Epoch 4 Batch 300 Loss0.3940

Epoch 4 Batch 400 Loss0.3909

Epoch 4 Batch 500 Loss0.4229

Epoch 4 Batch 600 Loss0.4785

Epoch 5 Loss 0.4288

Time taken for 1 epoch 62.66524386405945 sec

Epoch 5 Batch 0 Loss0.4234

Epoch 5 Batch 100 Loss0.4003

Epoch 5 Batch 200 Loss0.3822

Epoch 5 Batch 300 Loss0.4003

Epoch 5 Batch 400 Loss0.4401

Epoch 5 Batch 500 Loss0.4166

Epoch 5 Batch 600 Loss0.4121

Epoch 6 Loss 0.4105

Time taken for 1 epoch 62.635040044784546 sec

Epoch 6 Batch 0 Loss0.4217

Epoch 6 Batch 100 Loss0.3792

Epoch 6 Batch 200 Loss0.3546

Epoch 6 Batch 300 Loss0.4262

Epoch 6 Batch 400 Loss0.3972

Epoch 6 Batch 500 Loss0.4205

Epoch 6 Batch 600 Loss0.4106

Epoch 7 Loss 0.3948

Time taken for 1 epoch 62.20459198951721 sec

Epoch 7 Batch 0 Loss0.4029

Epoch 7 Batch 100 Loss0.3867

Epoch 7 Batch 200 Loss0.3561

Epoch 7 Batch 300 Loss0.4071

Epoch 7 Batch 400 Loss0.3518

Epoch 7 Batch 500 Loss0.3374

Epoch 7 Batch 600 Loss0.4123

Epoch 8 Loss 0.3800

Time taken for 1 epoch 62.43874192237854 sec

Epoch 8 Batch 0 Loss0.3988

Epoch 8 Batch 100 Loss0.3727

Epoch 8 Batch 200 Loss0.3404

Epoch 8 Batch 300 Loss0.4003

Epoch 8 Batch 400 Loss0.3585

Epoch 8 Batch 500 Loss0.3648

Epoch 8 Batch 600 Loss0.3708

Epoch 9 Loss 0.3669

Time taken for 1 epoch 62.57436680793762 sec

Epoch 9 Batch 0 Loss0.3254

Epoch 9 Batch 100 Loss0.3586

Epoch 9 Batch 200 Loss0.3896

Epoch 9 Batch 300 Loss0.3368

Epoch 9 Batch 400 Loss0.3545

Epoch 9 Batch 500 Loss0.3891

Epoch 9 Batch 600 Loss0.3517

Epoch 10 Loss 0.3542

Time taken for 1 epoch 62.56642198562622 sec

Epoch 10 Batch 0 Loss0.3382

Epoch 10 Batch 100 Loss0.3368

Epoch 10 Batch 200 Loss0.3446

Epoch 10 Batch 300 Loss0.3637

Epoch 10 Batch 400 Loss0.3380

Epoch 10 Batch 500 Loss0.3035

Epoch 10 Batch 600 Loss0.3626

Epoch 11 Loss 0.3421

Time taken for 1 epoch 64.20626091957092 sec

Epoch 11 Batch 0 Loss0.3474

Epoch 11 Batch 100 Loss0.2983

Epoch 11 Batch 200 Loss0.3731

Epoch 11 Batch 300 Loss0.3540

Epoch 11 Batch 400 Loss0.3853

Epoch 11 Batch 500 Loss0.2672

Epoch 11 Batch 600 Loss0.3430

Epoch 12 Loss 0.3306

Time taken for 1 epoch 62.776175022125244 sec

Epoch 12 Batch 0 Loss0.2609

Epoch 12 Batch 100 Loss0.2914

Epoch 12 Batch 200 Loss0.3136

Epoch 12 Batch 300 Loss0.3328

Epoch 12 Batch 400 Loss0.3439

Epoch 12 Batch 500 Loss0.3283

Epoch 12 Batch 600 Loss0.3042

Epoch 13 Loss 0.3197

Time taken for 1 epoch 61.50251746177673 sec

Epoch 13 Batch 0 Loss0.2715

Epoch 13 Batch 100 Loss0.3107

Epoch 13 Batch 200 Loss0.3293

Epoch 13 Batch 300 Loss0.2918

Epoch 13 Batch 400 Loss0.3100

Epoch 13 Batch 500 Loss0.3449

Epoch 13 Batch 600 Loss0.2889

Epoch 14 Loss 0.3083

Time taken for 1 epoch 64.8661618232727 sec

Epoch 14 Batch 0 Loss0.2541

Epoch 14 Batch 100 Loss0.3036

Epoch 14 Batch 200 Loss0.2791

Epoch 14 Batch 300 Loss0.3071

Epoch 14 Batch 400 Loss0.3154

Epoch 14 Batch 500 Loss0.2826

Epoch 14 Batch 600 Loss0.3369

Epoch 15 Loss 0.2990

Time taken for 1 epoch 62.334789752960205 sec

Epoch 15 Batch 0 Loss0.2757

Epoch 15 Batch 100 Loss0.2772

Epoch 15 Batch 200 Loss0.3063

Epoch 15 Batch 300 Loss0.3364

Epoch 15 Batch 400 Loss0.3521

Epoch 15 Batch 500 Loss0.3027

Epoch 15 Batch 600 Loss0.3091

Epoch 16 Loss 0.2899

Time taken for 1 epoch 64.35370945930481 sec

Epoch 16 Batch 0 Loss0.2646

Epoch 16 Batch 100 Loss0.2700

Epoch 16 Batch 200 Loss0.2696

Epoch 16 Batch 300 Loss0.2712

Epoch 16 Batch 400 Loss0.2898

Epoch 16 Batch 500 Loss0.2911

Epoch 16 Batch 600 Loss0.2862

Epoch 17 Loss 0.2804

Time taken for 1 epoch 63.70657563209534 sec

Epoch 17 Batch 0 Loss0.2584

Epoch 17 Batch 100 Loss0.2537

Epoch 17 Batch 200 Loss0.2527

Epoch 17 Batch 300 Loss0.2681

Epoch 17 Batch 400 Loss0.2571

Epoch 17 Batch 500 Loss0.2751

Epoch 17 Batch 600 Loss0.2569

Epoch 18 Loss 0.2719

Time taken for 1 epoch 63.51345372200012 sec

Epoch 18 Batch 0 Loss0.2715

Epoch 18 Batch 100 Loss0.2504

Epoch 18 Batch 200 Loss0.2251

Epoch 18 Batch 300 Loss0.2552

Epoch 18 Batch 400 Loss0.2892

Epoch 18 Batch 500 Loss0.2714

Epoch 18 Batch 600 Loss0.2447

Epoch 19 Loss 0.2650

Time taken for 1 epoch 63.62069797515869 sec

Epoch 19 Batch 0 Loss0.2748

Epoch 19 Batch 100 Loss0.2693

Epoch 19 Batch 200 Loss0.2552

Epoch 19 Batch 300 Loss0.2758

Epoch 19 Batch 400 Loss0.2714

Epoch 19 Batch 500 Loss0.2774

Epoch 19 Batch 600 Loss0.2807

Epoch 20 Loss 0.2583

Time taken for 1 epoch 63.734514474868774 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 64,

    GRU\_UNITS = 128,

    ATTN\_SIZE = 2,

)

Epoch 0 Batch 0 Loss1.3004

Epoch 0 Batch 100 Loss0.6190

Epoch 0 Batch 200 Loss0.6024

Epoch 0 Batch 300 Loss0.5410

Epoch 0 Batch 400 Loss0.5242

Epoch 0 Batch 500 Loss0.4701

Epoch 0 Batch 600 Loss0.4336

Epoch 1 Loss 0.5916

Time taken for 1 epoch 116.26714563369751 sec

Epoch 1 Batch 0 Loss0.5190

Epoch 1 Batch 100 Loss0.5095

Epoch 1 Batch 200 Loss0.4527

Epoch 1 Batch 300 Loss0.5153

Epoch 1 Batch 400 Loss0.3931

Epoch 1 Batch 500 Loss0.4288

Epoch 1 Batch 600 Loss0.4109

Epoch 2 Loss 0.4492

Time taken for 1 epoch 66.68127942085266 sec

Epoch 2 Batch 0 Loss0.3864

Epoch 2 Batch 100 Loss0.4014

Epoch 2 Batch 200 Loss0.3855

Epoch 2 Batch 300 Loss0.3829

Epoch 2 Batch 400 Loss0.3968

Epoch 2 Batch 500 Loss0.4461

Epoch 2 Batch 600 Loss0.3770

Epoch 3 Loss 0.3859

Time taken for 1 epoch 66.85419344902039 sec

Epoch 3 Batch 0 Loss0.3441

Epoch 3 Batch 100 Loss0.3038

Epoch 3 Batch 200 Loss0.3504

Epoch 3 Batch 300 Loss0.4013

Epoch 3 Batch 400 Loss0.3716

Epoch 3 Batch 500 Loss0.3286

Epoch 3 Batch 600 Loss0.3177

Epoch 4 Loss 0.3437

Time taken for 1 epoch 66.26851058006287 sec

Epoch 4 Batch 0 Loss0.2935

Epoch 4 Batch 100 Loss0.3165

Epoch 4 Batch 200 Loss0.3184

Epoch 4 Batch 300 Loss0.3548

Epoch 4 Batch 400 Loss0.2776

Epoch 4 Batch 500 Loss0.3179

Epoch 4 Batch 600 Loss0.3099

Epoch 5 Loss 0.3134

Time taken for 1 epoch 66.98063492774963 sec

Epoch 5 Batch 0 Loss0.3309

Epoch 5 Batch 100 Loss0.2938

Epoch 5 Batch 200 Loss0.2883

Epoch 5 Batch 300 Loss0.2681

Epoch 5 Batch 400 Loss0.3415

Epoch 5 Batch 500 Loss0.3008

Epoch 5 Batch 600 Loss0.2896

Epoch 6 Loss 0.2998

Time taken for 1 epoch 65.74953150749207 sec

Epoch 6 Batch 0 Loss0.2450

Epoch 6 Batch 100 Loss0.2817

Epoch 6 Batch 200 Loss0.2525

Epoch 6 Batch 300 Loss0.2318

Epoch 6 Batch 400 Loss0.2590

Epoch 6 Batch 500 Loss0.2819

Epoch 6 Batch 600 Loss0.2436

Epoch 7 Loss 0.2825

Time taken for 1 epoch 66.30567717552185 sec

Epoch 7 Batch 0 Loss0.2836

Epoch 7 Batch 100 Loss0.2541

Epoch 7 Batch 200 Loss0.2880

Epoch 7 Batch 300 Loss0.2803

Epoch 7 Batch 400 Loss0.2579

Epoch 7 Batch 500 Loss0.2694

Epoch 7 Batch 600 Loss0.3012

Epoch 8 Loss 0.2760

Time taken for 1 epoch 66.1033501625061 sec

Epoch 8 Batch 0 Loss0.2226

Epoch 8 Batch 100 Loss0.3188

Epoch 8 Batch 200 Loss0.2997

Epoch 8 Batch 300 Loss0.2643

Epoch 8 Batch 400 Loss0.2720

Epoch 8 Batch 500 Loss0.2879

Epoch 8 Batch 600 Loss0.3160

Epoch 9 Loss 0.2679

Time taken for 1 epoch 67.24814105033875 sec

Epoch 9 Batch 0 Loss0.2558

Epoch 9 Batch 100 Loss0.2405

Epoch 9 Batch 200 Loss0.2350

Epoch 9 Batch 300 Loss0.2845

Epoch 9 Batch 400 Loss0.2725

Epoch 9 Batch 500 Loss0.2413

Epoch 9 Batch 600 Loss0.2601

Epoch 10 Loss 0.2520

Time taken for 1 epoch 66.70499444007874 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 32,

    GRU\_UNITS = 256,

    ATTN\_SIZE = 2,

)

Epoch 0 Batch 0 Loss1.2733

Epoch 0 Batch 100 Loss0.6647

Epoch 0 Batch 200 Loss0.6354

Epoch 0 Batch 300 Loss0.6463

Epoch 0 Batch 400 Loss0.5721

Epoch 0 Batch 500 Loss0.5471

Epoch 0 Batch 600 Loss0.5072

Epoch 1 Loss 0.6123

Time taken for 1 epoch 111.86861848831177 sec

Epoch 1 Batch 0 Loss0.5487

Epoch 1 Batch 100 Loss0.5062

Epoch 1 Batch 200 Loss0.4678

Epoch 1 Batch 300 Loss0.4994

Epoch 1 Batch 400 Loss0.4500

Epoch 1 Batch 500 Loss0.5053

Epoch 1 Batch 600 Loss0.4701

Epoch 2 Loss 0.4696

Time taken for 1 epoch 63.16306257247925 sec

Epoch 2 Batch 0 Loss0.4762

Epoch 2 Batch 100 Loss0.4354

Epoch 2 Batch 200 Loss0.4059

Epoch 2 Batch 300 Loss0.3858

Epoch 2 Batch 400 Loss0.3916

Epoch 2 Batch 500 Loss0.4208

Epoch 2 Batch 600 Loss0.3605

Epoch 3 Loss 0.4099

Time taken for 1 epoch 62.62613320350647 sec

Epoch 3 Batch 0 Loss0.3938

Epoch 3 Batch 100 Loss0.3683

Epoch 3 Batch 200 Loss0.3655

Epoch 3 Batch 300 Loss0.3748

Epoch 3 Batch 400 Loss0.3224

Epoch 3 Batch 500 Loss0.3666

Epoch 3 Batch 600 Loss0.4219

Epoch 4 Loss 0.3641

Time taken for 1 epoch 62.93504738807678 sec

Epoch 4 Batch 0 Loss0.3380

Epoch 4 Batch 100 Loss0.3673

Epoch 4 Batch 200 Loss0.3003

Epoch 4 Batch 300 Loss0.2755

Epoch 4 Batch 400 Loss0.3242

Epoch 4 Batch 500 Loss0.3556

Epoch 4 Batch 600 Loss0.3444

Epoch 5 Loss 0.3220

Time taken for 1 epoch 62.99858379364014 sec

Epoch 5 Batch 0 Loss0.2964

Epoch 5 Batch 100 Loss0.2734

Epoch 5 Batch 200 Loss0.3039

Epoch 5 Batch 300 Loss0.3203

Epoch 5 Batch 400 Loss0.2672

Epoch 5 Batch 500 Loss0.2927

Epoch 5 Batch 600 Loss0.2704

Epoch 6 Loss 0.2849

Time taken for 1 epoch 63.0986545085907 sec

Epoch 6 Batch 0 Loss0.2697

Epoch 6 Batch 100 Loss0.2404

Epoch 6 Batch 200 Loss0.2917

Epoch 6 Batch 300 Loss0.2121

Epoch 6 Batch 400 Loss0.2290

Epoch 6 Batch 500 Loss0.2747

Epoch 6 Batch 600 Loss0.2691

Epoch 7 Loss 0.2523

Time taken for 1 epoch 63.025723695755005 sec

Epoch 7 Batch 0 Loss0.2388

Epoch 7 Batch 100 Loss0.2006

Epoch 7 Batch 200 Loss0.2292

Epoch 7 Batch 300 Loss0.2585

Epoch 7 Batch 400 Loss0.2076

Epoch 7 Batch 500 Loss0.2986

Epoch 7 Batch 600 Loss0.2361

Epoch 8 Loss 0.2224

Time taken for 1 epoch 62.74348998069763 sec

Epoch 8 Batch 0 Loss0.1622

Epoch 8 Batch 100 Loss0.1870

Epoch 8 Batch 200 Loss0.1688

Epoch 8 Batch 300 Loss0.2273

Epoch 8 Batch 400 Loss0.1898

Epoch 8 Batch 500 Loss0.2377

Epoch 8 Batch 600 Loss0.2067

Epoch 9 Loss 0.1973

Time taken for 1 epoch 62.62703824043274 sec

Epoch 9 Batch 0 Loss0.1807

Epoch 9 Batch 100 Loss0.1952

Epoch 9 Batch 200 Loss0.1781

Epoch 9 Batch 300 Loss0.1766

Epoch 9 Batch 400 Loss0.2032

Epoch 9 Batch 500 Loss0.1878

Epoch 9 Batch 600 Loss0.1772

Epoch 10 Loss 0.1770

Time taken for 1 epoch 63.42581057548523 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 16,

    GRU\_UNITS = 256,

    ATTN\_SIZE = 2,

)

Epoch 0 Batch 0 Loss1.3058

Epoch 0 Batch 100 Loss0.7570

Epoch 0 Batch 200 Loss0.6468

Epoch 0 Batch 300 Loss0.5528

Epoch 0 Batch 400 Loss0.5125

Epoch 0 Batch 500 Loss0.5161

Epoch 0 Batch 600 Loss0.4997

Epoch 1 Loss 0.6120

Time taken for 1 epoch 109.6115882396698 sec

Epoch 1 Batch 0 Loss0.5010

Epoch 1 Batch 100 Loss0.4519

Epoch 1 Batch 200 Loss0.5220

Epoch 1 Batch 300 Loss0.4399

Epoch 1 Batch 400 Loss0.4007

Epoch 1 Batch 500 Loss0.4339

Epoch 1 Batch 600 Loss0.3884

Epoch 2 Loss 0.4485

Time taken for 1 epoch 60.467719316482544 sec

Epoch 2 Batch 0 Loss0.4161

Epoch 2 Batch 100 Loss0.3958

Epoch 2 Batch 200 Loss0.3974

Epoch 2 Batch 300 Loss0.3404

Epoch 2 Batch 400 Loss0.3621

Epoch 2 Batch 500 Loss0.3516

Epoch 2 Batch 600 Loss0.3720

Epoch 3 Loss 0.3702

Time taken for 1 epoch 61.35898494720459 sec

Epoch 3 Batch 0 Loss0.2889

Epoch 3 Batch 100 Loss0.3207

Epoch 3 Batch 200 Loss0.2956

Epoch 3 Batch 300 Loss0.2924

Epoch 3 Batch 400 Loss0.2713

Epoch 3 Batch 500 Loss0.2607

Epoch 3 Batch 600 Loss0.2926

Epoch 4 Loss 0.3097

Time taken for 1 epoch 61.44639849662781 sec

Epoch 4 Batch 0 Loss0.3068

Epoch 4 Batch 100 Loss0.2679

Epoch 4 Batch 200 Loss0.2171

Epoch 4 Batch 300 Loss0.2730

Epoch 4 Batch 400 Loss0.2044

Epoch 4 Batch 500 Loss0.2753

Epoch 4 Batch 600 Loss0.2149

Epoch 5 Loss 0.2599

Time taken for 1 epoch 60.983553409576416 sec

Epoch 5 Batch 0 Loss0.2054

Epoch 5 Batch 100 Loss0.2044

Epoch 5 Batch 200 Loss0.1988

Epoch 5 Batch 300 Loss0.2308

Epoch 5 Batch 400 Loss0.1807

Epoch 5 Batch 500 Loss0.2127

Epoch 5 Batch 600 Loss0.1852

Epoch 6 Loss 0.2203

Time taken for 1 epoch 61.46408462524414 sec

Epoch 6 Batch 0 Loss0.1894

Epoch 6 Batch 100 Loss0.1647

Epoch 6 Batch 200 Loss0.2449

Epoch 6 Batch 300 Loss0.1958

Epoch 6 Batch 400 Loss0.2157

Epoch 6 Batch 500 Loss0.1688

Epoch 6 Batch 600 Loss0.1692

Epoch 7 Loss 0.1886

Time taken for 1 epoch 61.743051290512085 sec

Epoch 7 Batch 0 Loss0.1779

Epoch 7 Batch 100 Loss0.1455

Epoch 7 Batch 200 Loss0.1417

Epoch 7 Batch 300 Loss0.1872

Epoch 7 Batch 400 Loss0.1563

Epoch 7 Batch 500 Loss0.2035

Epoch 7 Batch 600 Loss0.1804

Epoch 8 Loss 0.1647

Time taken for 1 epoch 61.57805895805359 sec

Epoch 8 Batch 0 Loss0.1490

Epoch 8 Batch 100 Loss0.1186

Epoch 8 Batch 200 Loss0.1639

Epoch 8 Batch 300 Loss0.1580

Epoch 8 Batch 400 Loss0.1397

Epoch 8 Batch 500 Loss0.1428

Epoch 8 Batch 600 Loss0.2168

Epoch 9 Loss 0.1479

Time taken for 1 epoch 61.52410626411438 sec

Epoch 9 Batch 0 Loss0.1202

Epoch 9 Batch 100 Loss0.1195

Epoch 9 Batch 200 Loss0.1248

Epoch 9 Batch 300 Loss0.1404

Epoch 9 Batch 400 Loss0.1516

Epoch 9 Batch 500 Loss0.1682

Epoch 9 Batch 600 Loss0.1457

Epoch 10 Loss 0.1343

Time taken for 1 epoch 61.476850509643555 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 8,

    GRU\_UNITS = 256,

    ATTN\_SIZE = 2,

)

Epoch 0 Batch 0 Loss1.4220

Epoch 0 Batch 100 Loss0.7225

Epoch 0 Batch 200 Loss0.6639

Epoch 0 Batch 300 Loss0.6065

Epoch 0 Batch 400 Loss0.5575

Epoch 0 Batch 500 Loss0.5197

Epoch 0 Batch 600 Loss0.4037

Epoch 1 Loss 0.6084

Time taken for 1 epoch 107.19780564308167 sec

Epoch 1 Batch 0 Loss0.4920

Epoch 1 Batch 100 Loss0.4329

Epoch 1 Batch 200 Loss0.4434

Epoch 1 Batch 300 Loss0.4219

Epoch 1 Batch 400 Loss0.4555

Epoch 1 Batch 500 Loss0.4219

Epoch 1 Batch 600 Loss0.3953

Epoch 2 Loss 0.4449

Time taken for 1 epoch 60.890280961990356 sec

Epoch 2 Batch 0 Loss0.4070

Epoch 2 Batch 100 Loss0.3332

Epoch 2 Batch 200 Loss0.3387

Epoch 2 Batch 300 Loss0.3890

Epoch 2 Batch 400 Loss0.3816

Epoch 2 Batch 500 Loss0.3794

Epoch 2 Batch 600 Loss0.2836

Epoch 3 Loss 0.3576

Time taken for 1 epoch 60.660072326660156 sec

Epoch 3 Batch 0 Loss0.2969

Epoch 3 Batch 100 Loss0.3059

Epoch 3 Batch 200 Loss0.2975

Epoch 3 Batch 300 Loss0.2683

Epoch 3 Batch 400 Loss0.3298

Epoch 3 Batch 500 Loss0.2860

Epoch 3 Batch 600 Loss0.2540

Epoch 4 Loss 0.2893

Time taken for 1 epoch 61.12022566795349 sec

Epoch 4 Batch 0 Loss0.2599

Epoch 4 Batch 100 Loss0.2260

Epoch 4 Batch 200 Loss0.2308

Epoch 4 Batch 300 Loss0.2151

Epoch 4 Batch 400 Loss0.2768

Epoch 4 Batch 500 Loss0.2418

Epoch 4 Batch 600 Loss0.2175

Epoch 5 Loss 0.2403

Time taken for 1 epoch 61.892218828201294 sec

Epoch 5 Batch 0 Loss0.1647

Epoch 5 Batch 100 Loss0.2034

Epoch 5 Batch 200 Loss0.1957

Epoch 5 Batch 300 Loss0.2299

Epoch 5 Batch 400 Loss0.1880

Epoch 5 Batch 500 Loss0.1869

Epoch 5 Batch 600 Loss0.2156

Epoch 6 Loss 0.2054

Time taken for 1 epoch 60.9693284034729 sec

Epoch 6 Batch 0 Loss0.1969

Epoch 6 Batch 100 Loss0.1420

Epoch 6 Batch 200 Loss0.1611

Epoch 6 Batch 300 Loss0.1567

Epoch 6 Batch 400 Loss0.1559

Epoch 6 Batch 500 Loss0.1531

Epoch 6 Batch 600 Loss0.2037

Epoch 7 Loss 0.1789

Time taken for 1 epoch 61.2753221988678 sec

Epoch 7 Batch 0 Loss0.1262

Epoch 7 Batch 100 Loss0.1433

Epoch 7 Batch 200 Loss0.1298

Epoch 7 Batch 300 Loss0.1396

Epoch 7 Batch 400 Loss0.1565

Epoch 7 Batch 500 Loss0.1685

Epoch 7 Batch 600 Loss0.1708

Epoch 8 Loss 0.1593

Time taken for 1 epoch 61.7517728805542 sec

Epoch 8 Batch 0 Loss0.1369

Epoch 8 Batch 100 Loss0.1269

Epoch 8 Batch 200 Loss0.1519

Epoch 8 Batch 300 Loss0.1438

Epoch 8 Batch 400 Loss0.1394

Epoch 8 Batch 500 Loss0.1300

Epoch 8 Batch 600 Loss0.1407

Epoch 9 Loss 0.1434

Time taken for 1 epoch 61.65241050720215 sec

Epoch 9 Batch 0 Loss0.1092

Epoch 9 Batch 100 Loss0.1377

Epoch 9 Batch 200 Loss0.1485

Epoch 9 Batch 300 Loss0.0992

Epoch 9 Batch 400 Loss0.1461

Epoch 9 Batch 500 Loss0.1201

Epoch 9 Batch 600 Loss0.1396

Epoch 10 Loss 0.1324

Time taken for 1 epoch 61.45880889892578 sec

model\_hype = Namespace(

    EMBEDDING\_SIZE = 4,

    GRU\_UNITS = 256,

    ATTN\_SIZE = 2,

)

Epoch 0 Batch 0 Loss1.2841

Epoch 0 Batch 100 Loss0.6290

Epoch 0 Batch 200 Loss0.6177

Epoch 0 Batch 300 Loss0.5626

Epoch 0 Batch 400 Loss0.5400

Epoch 0 Batch 500 Loss0.5498

Epoch 0 Batch 600 Loss0.4392

Epoch 1 Loss 0.6050

Time taken for 1 epoch 107.50903177261353 sec

Epoch 1 Batch 0 Loss0.4405

Epoch 1 Batch 100 Loss0.4774

Epoch 1 Batch 200 Loss0.4060

Epoch 1 Batch 300 Loss0.3935

Epoch 1 Batch 400 Loss0.4114

Epoch 1 Batch 500 Loss0.3900

Epoch 1 Batch 600 Loss0.3226

Epoch 2 Loss 0.4103

Time taken for 1 epoch 60.095754623413086 sec

Epoch 2 Batch 0 Loss0.3585

Epoch 2 Batch 100 Loss0.3069

Epoch 2 Batch 200 Loss0.3282

Epoch 2 Batch 300 Loss0.3214

Epoch 2 Batch 400 Loss0.3678

Epoch 2 Batch 500 Loss0.3799

Epoch 2 Batch 600 Loss0.2810

Epoch 3 Loss 0.3252

Time taken for 1 epoch 59.77031397819519 sec

Epoch 3 Batch 0 Loss0.2522

Epoch 3 Batch 100 Loss0.2706

Epoch 3 Batch 200 Loss0.2543

Epoch 3 Batch 300 Loss0.2535

Epoch 3 Batch 400 Loss0.2760

Epoch 3 Batch 500 Loss0.2551

Epoch 3 Batch 600 Loss0.2492

Epoch 4 Loss 0.2704

Time taken for 1 epoch 60.41859221458435 sec

Epoch 4 Batch 0 Loss0.2416

Epoch 4 Batch 100 Loss0.2620

Epoch 4 Batch 200 Loss0.2026

Epoch 4 Batch 300 Loss0.2291

Epoch 4 Batch 400 Loss0.2225

Epoch 4 Batch 500 Loss0.2033

Epoch 4 Batch 600 Loss0.2408

Epoch 5 Loss 0.2319

Time taken for 1 epoch 59.77333211898804 sec

Epoch 5 Batch 0 Loss0.1862

Epoch 5 Batch 100 Loss0.1504

Epoch 5 Batch 200 Loss0.1968

Epoch 5 Batch 300 Loss0.2278

Epoch 5 Batch 400 Loss0.2374

Epoch 5 Batch 500 Loss0.2285

Epoch 5 Batch 600 Loss0.2138

Epoch 6 Loss 0.2028

Time taken for 1 epoch 60.88700723648071 sec

Epoch 6 Batch 0 Loss0.1512

Epoch 6 Batch 100 Loss0.1820

Epoch 6 Batch 200 Loss0.1866

Epoch 6 Batch 300 Loss0.1733

Epoch 6 Batch 400 Loss0.1812

Epoch 6 Batch 500 Loss0.1675

Epoch 6 Batch 600 Loss0.2069

Epoch 7 Loss 0.1813

Time taken for 1 epoch 61.077032804489136 sec

Epoch 7 Batch 0 Loss0.1697

Epoch 7 Batch 100 Loss0.1234

Epoch 7 Batch 200 Loss0.1803

Epoch 7 Batch 300 Loss0.1526

Epoch 7 Batch 400 Loss0.1631

Epoch 7 Batch 500 Loss0.1865

Epoch 7 Batch 600 Loss0.1631

Epoch 8 Loss 0.1650

Time taken for 1 epoch 61.13740634918213 sec

Epoch 8 Batch 0 Loss0.1468

Epoch 8 Batch 100 Loss0.1672

Epoch 8 Batch 200 Loss0.1414

Epoch 8 Batch 300 Loss0.1208

Epoch 8 Batch 400 Loss0.1939

Epoch 8 Batch 500 Loss0.1516

Epoch 8 Batch 600 Loss0.1578

Epoch 9 Loss 0.1512

Time taken for 1 epoch 61.51504158973694 sec

Epoch 9 Batch 0 Loss0.1431

Epoch 9 Batch 100 Loss0.1276

Epoch 9 Batch 200 Loss0.1146

Epoch 9 Batch 300 Loss0.1159

Epoch 9 Batch 400 Loss0.1271

Epoch 9 Batch 500 Loss0.1383

Epoch 9 Batch 600 Loss0.1728

Epoch 10 Loss 0.1401

Time taken for 1 epoch 61.70148944854736 sec