Q1:
Fine–tuned mT5–small model, with num_heads=6, num_layers = 8, and num_decoder_layers =8.
It add "summarize:" as the prefix of the input text and then generate the result.

For preprocess part, I add "summarize:" as the prefix of input text, truncate the input to at most 1024 long, title to 128 long, and use mT5 tokenizer to tokenize the inputs.

Q2:
Learning rate = 5e–05
per_device_train_batch_size=4
training epoch = 4

The learning rate and batch size is the default value, I first train for 3 epoch and did not pass the baseline, so I train it for 1 more epoch.

Q3:

Greedy: always choose the most probable next word.

Beam search: Track a number of the most probable sequences, and choose the most probable one among them

Top–K sampling = Randomly choose a word among k of the most probable next word.

Top–p Sampling = Randomly choose a word among the most probable next words such that their probability have sum equals to p.

Temperature = A softmax function to make the probability distribution of the next word more extreme.

Comparing the result, I choose beam_search, beam_num = 6 as my final generation strategy.

Greedy:

```
{
  "rouge-1": {
    "r": 0.2126501811088697,
    "p": 0.2763711859240302,
    "f": 0.23023735865482398
  },
  "rouge-2": {
    "r": 0.08097073279231462,
    "p": 0.09913806680476372,
    "f": 0.08544245264750118
  },
  "rouge-l": {
    "r": 0.19184843044075336,
    "p": 0.2498902264636683,
    "f": 0.20776273835620496
  }
}
```

beam, beam=3:

```
{
  "rouge-1": {
    "r": 0.23339813129096992,
    "p": 0.2769154150727121,
    "f": 0.24311591316805228
  },
  "rouge-2": {
    "r": 0.0939836191448558,
    "p": 0.10815524179945551,
    "f": 0.09660886287684785
  },
  "rouge-l": {
    "r": 0.20970248849662143,
    "p": 0.24988899735021847,
    "f": 0.21872167826545402
  }
}
```

beam, beam=6:

```
{
  "rouge-1": {
    "r": 0.23859464079840556,
    "p": 0.27230263355729495,
    "f": 0.24439640005447166
```

```
    },
    "rouge-2": {
      "r": 0.0974792130916469,
      "p": 0.10791490836675656,
      "f": 0.09836168325484893
    },
    "rouge-l": {
      "r": 0.21407345211377146,
      "p": 0.24509606768774184,
      "f": 0.21944137190343282
    }
}
top_k, k=20:
{
    "rouge-1": {
      "r": 0.1870152250590569,
      "p": 0.2154084821739401,
      "f": 0.19304345838743098
    },
    "rouge-2": {
      "r": 0.06121567164860382,
      "p": 0.06715420385402084,
      "f": 0.061495569696759074
    },
    "rouge-l": {
      "r": 0.16568124621757693,
      "p": 0.1906115802092193,
      "f": 0.1708050365969121
    }
}
top_k, k=10:
{
    "rouge-1": {
      "r": 0.19633383963037113,
      "p": 0.23121665178145193,
      "f": 0.20480710745886582
    },
    "rouge-2": {
      "r": 0.066570236859016,
      "p": 0.07432254734807522,
      "f": 0.06756644403522548
    },
    "rouge-l": {
      "r": 0.17405227761377703,
      "p": 0.204998261430753,
      "f": 0.1814607934707028
    }
}

temp, temp=70:
{
    "rouge-1": {
      "r": 0.1841484387391865,
      "p": 0.21634939792057736,
      "f": 0.19199611511491754
```

```
  },
  "rouge-2": {
    "r": 0.06354783076528567,
    "p": 0.071133223873776,
    "f": 0.06461639011611278
  },
  "rouge-l": {
    "r": 0.16514033621214308,
    "p": 0.19395439611843793,
    "f": 0.1720760719257607
  }
}

temp, temp=80:
{
  "rouge-1": {
    "r": 0.15793804285972318,
    "p": 0.17683961402365878,
    "f": 0.1612093395189474
  },
  "rouge-2": {
    "r": 0.05011690745420843,
    "p": 0.05420440262822638,
    "f": 0.0501200802802458
  },
  "rouge-l": {
    "r": 0.14169759398320825,
    "p": 0.15873191306000736,
    "f": 0.14459668887321167
  }
}
top_p, p=0.85:
{
  "rouge-1": {
    "r": 0.14868989298058113,
    "p": 0.16270481021848623,
    "f": 0.1501229295731166
  },
  "rouge-2": {
    "r": 0.04660235752149222,
    "p": 0.04937745655927857,
    "f": 0.04609732085039801
  },
  "rouge-l": {
    "r": 0.13274712489219026,
    "p": 0.1454152818846778,
    "f": 0.13398705529497046
  }
}
top_p, p=0.95:
{
  "rouge-1": {
    "r": 0.1361547973292333,
    "p": 0.14460117976605757,
    "f": 0.1353773307734829
```

```
  },
  "rouge-2": {
    "r": 0.04177444971485365,
    "p": 0.04244103873328997,
    "f": 0.04048901386511494
  },
  "rouge-l": {
    "r": 0.12233540219401458,
    "p": 0.1301910934401233,
    "f": 0.12166342451520816
  }
}
```