Large language models: The GPT architecture

Steps in making a LLM

01 Language processing

02 Defining a models architecture

03 Model training and evaluation

01

Language processing

Text cleanup

XXVI. "But, O noble sir," etc. Plato, Gorgias, 512 D. XXVII. "And as for those parts," etc. A quotation from Euripides, Chryssipus, frag. 839 (Nauck).
"With meats," etc. From Euripides, Supplices, 1110. XXXIII. "They both," i.e. life and wrestling.
"Says he" (63): Plato, quoted by Epictetus, Arr. i. 28, 2 and 22.
XXXVII. "How know we," etc. The Greek means: "how know we whether Telauges were not nobler in character than Sophocles?" The allusion is unknown.
XXVII. "Frost" The word is written by Casaubon as a proper name, "Pagus.'
"The hardihood of Socrates was famous"; see Plato, Siymposium, p. 220.


```
Words that might be misleading for our model:  ['"a', '"as', '"do', '"en', '"i', '"it', '"m.', '"my', '"or', '"to', '&c.', "'a", "'as", "'be", "'do", "'he", "'i", "'if", "'in", "'it", "'m.", "'my", "'no", "'o", "'o,", "'of", "'to", "'we", '(1)', '(2)', '(3)', '(4)', '(5)', '(6)', '(7)', '(8)', '(a', '(as', '(be', '(c.', '(he', '(if', '(in', '(is', '(of', '(or', '(to', '(v.', '1', '1,', '1.', '10.', '11', '12', '12.', '14', '140', '143', '145', '169', '17', '17,', '174', '19', '19.', '2', '2.', '21.', '22,', '22.', '250', '26,', '28', '28,', '287', '2nd', '3,', '3.', '31', '345', '350', '353', '37.', '373', '4.', '412', '44.', '46', '486', '4th', '5.', '512', '5th', '6,', '6.', '66', '6th', '7', '7,', '7.', '757', '79', '8.', '839', '898', '9th', '[1]', '[2]', '[3]', '[4]', '[5]', '[6]', '[7]', '[8]', '[9]', '_in', 'a', 'a.', 'a.,', 'act', 'ad', 'ad.', 'add', 'ado', 'age', 'ago', 'ah,', 'aid', 'aim', 'air', 'all', 'am', 'am,', 'an', 'and', 'any', 'apt', 'are', 'arm', 'art', 'as', 'as,', 'ask', 'at', 'at,', 'at.', 'at?', 'axe', 'b.', 'b.c', 'bad', 'bag', 'be', 'be)', 'be,', 'be.', 'be:', 'be;', 'be?', 'bed', 'bee', 'bet', 'bid', 'big', 'bit', 'boy', 'bud', 'but', 'by', 'by,', 'by;', 'c.', 'can', 'cap', 'cry', 'cut', 'd.', 'day', 'de', 'did', 'die', 'dig', 'dio', 'do', 'do!', 'do,', 'do.', 'do:', 'do;', 'do?', 'dog', 'don', 'dry', 'due', 'dye', 'ear', 'eat', 'ed.', 'end', 'ep.', 'et', 'eye', 'f.', 'far', 'few', 'fig', 'fit', 'fix', 'fl.', 'fly', 'foe', 'for', 'fro', 'g.', 'gar', 'get', 'go', 'go,', 'god', 'got', 'gr.', 'h.', 'had', 'hap', 'has', 'he', 'he"', 'he)', 'he,', 'he.', 'he:', 'her', 'hid', 'him', 'his', 'hit', 'ho!', 'ho,', 'hot', 'how', 'hut', 'i', 'i)', 'i,', 'i.', 'i.,', 'i?', 'if', 'if,', 'ii', 'ii,', 'ii.', 'ill', 'imp', 'in', 'in,', 'in.', 'in;', 'in?', 'ir', 'is', 'is!', 'is)', 'is,', 'is.', 'is:', 'is;', 'it', 'it!', 'it)', 'it,', 'it.', 'it:', 'it;', 'it?', 'its', 'iv', 'iv.', 'ix.', 'j.', 'joy', 'kl', 'l.', 'law', 'lay', 'led', 'leg', 'let', 'li.', 'lie', 'lot', 'low', 'lv.', 'm', 'm,', 'm.', 'mad', 'mai', 'man', 'may', 'md', 'me', 'me!', 'me,', 'me.', 'me:', 'me;', 'me?', 'men', 'met', 'mh', 'mr.', 'ms.', 'my', 'n.', 'nay', 'net', 'new', 'no', 'no,', 'no.', 'no?', 'nor', 'not', 'now', 'o', 'odd', 'of', 'of,', 'of.', 'of:', 'of;', 'of?', 'off', 'oft', 'oh,', 'oil', 'old', 'on', 'on,', 'on.', 'on;', 'one', 'op.', 'or', 'or,', 'our', 'out', 'owe', 'own', 'ox,', 'p.', 'par', 'pay', 'pen', 'pig', 'put', 'q.', 'r.', 'ram', 'rid', 'rob', 'rod', 'run', 's.', 'sap', 'saw', 'say', 'sc.', 'see', 'set', 'she', 'sin', 'sit', 'six', 'sky', 'so', 'so)', 'so,', 'so.', 'so:', 'so;', 'so?', 'son', 'st.', 'sum', 'sun', 'ta', 'ten', 'the', 'thy', 'ti,', 'tin', 'tit', 'to', 'to,', 'too', 'top', 'try', 'two', 'up', 'up,', 'up.', 'us', 'us,', 'us.', 'us:', 'us;', 'us?', 'use', 'v', 'v.', 'vex', 'vi', 'vi.', 'vii', 'vow', 'war', 'was', 'wax', 'way', 'we', 'we:', 'wet', 'who', 'why', 'win', 'wit', 'woe', 'won', 'x', 'x.', 'xi', 'xi.', 'xii', 'xiv', 'xl.', 'xv.', 'xx.', 'ye', 'yea', 'yet', 'you', '§', 'μή,', 'μὲ.', 'τὸ', 'τῇ']
```

Tokens?

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

Tokens?

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

Words:

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

Tokens?

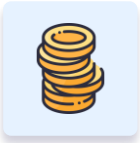
I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

Words:

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

Letters:

I. A man must not only

Tokens? 

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

Words:

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

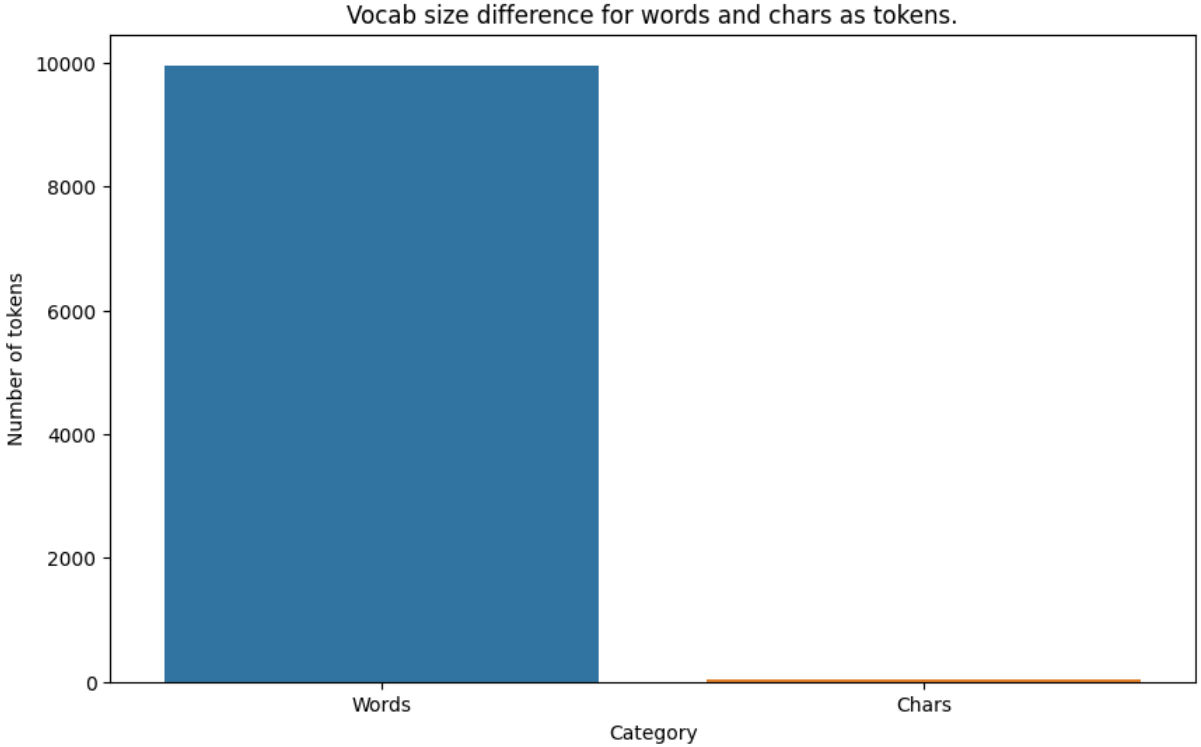
Letters:

I. A man must not only

Subwords:

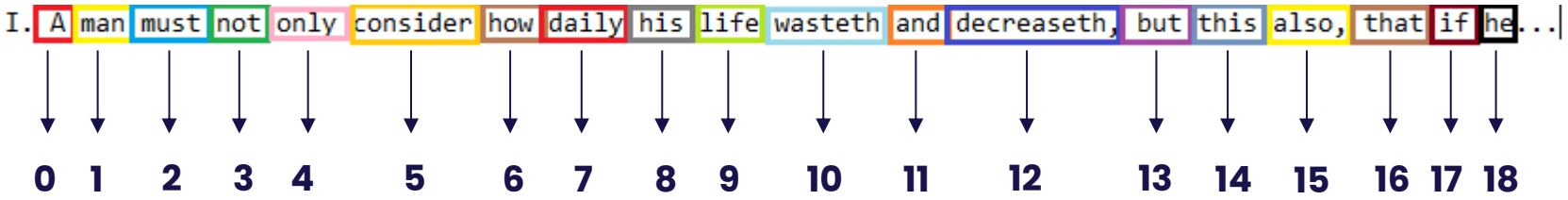
I. A man must not only consider how daily his life wasteth and decreaseth,

Tokens?

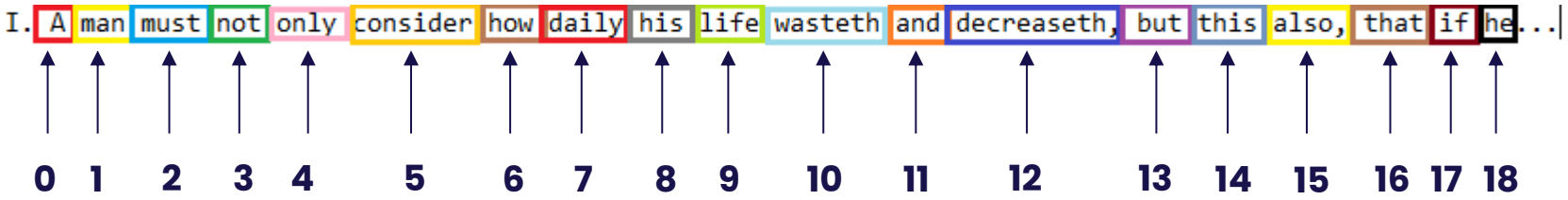

Vocab size difference for words and chars as tokens.
Number of tokens
10000
8000
6000
4000
2000
0
Words
Chars
Category

Encoding

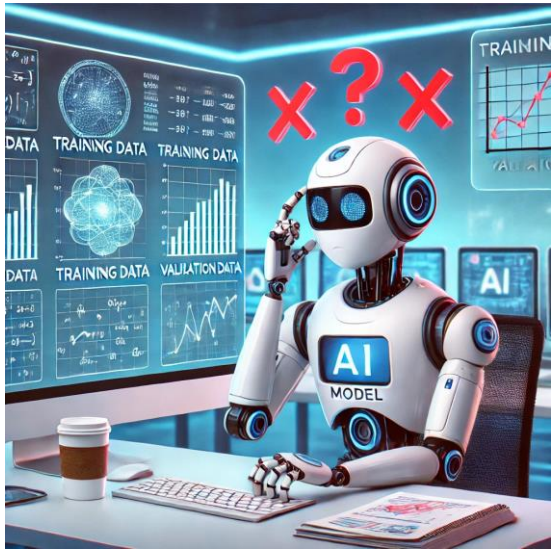
Encoding


I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

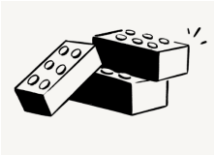
Decoding


I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Training and validation sets

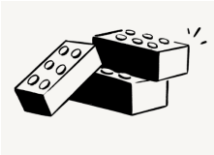


“Block”



I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

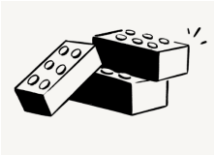
“Block”



“Block” of size 8 tokens

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

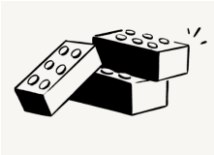
“Block”



“Block” of size 8 tokens

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

“Block”

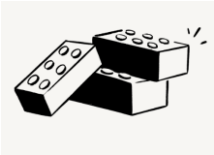


“Block” of size 8 tokens

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

“Block”



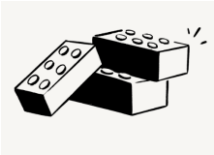
“Block” of size 8 tokens

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

“Block”



“Block” of size 8 tokens

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

“Batch”

“Batch” of size 4 “Blocks”

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

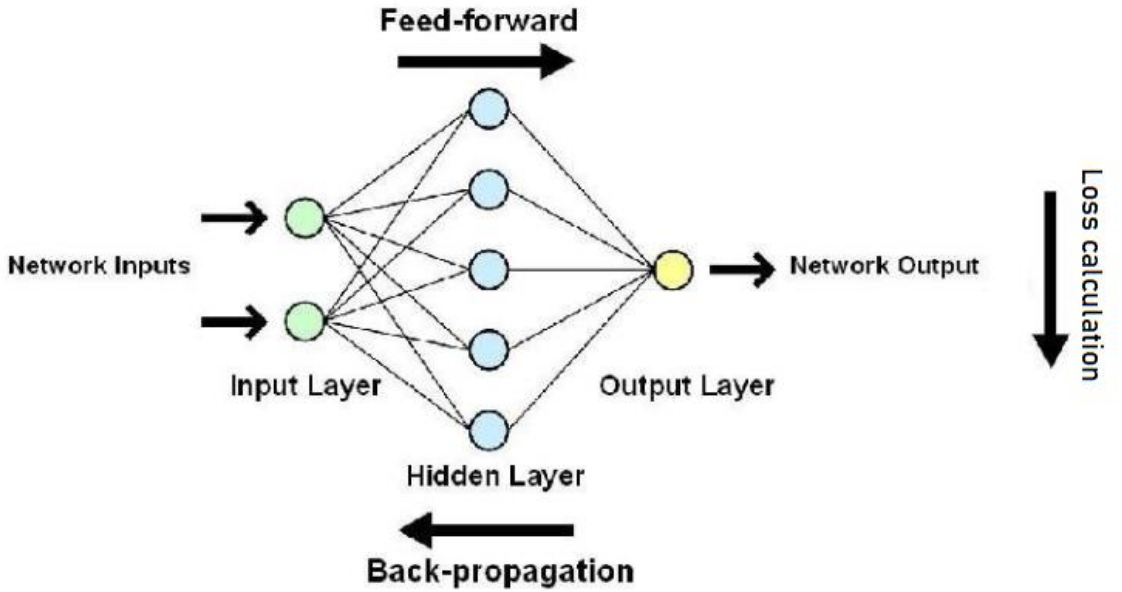
I. A man must not only consider how daily his life wasteth and decreaseth, but this also, that if he...

PCI Express 3.0 Host Interface
GigaThread Engine
GPC
TPC
SM
Memory Controller
HBM2
L2 Cache
High-Speed Hub
NVLink

03

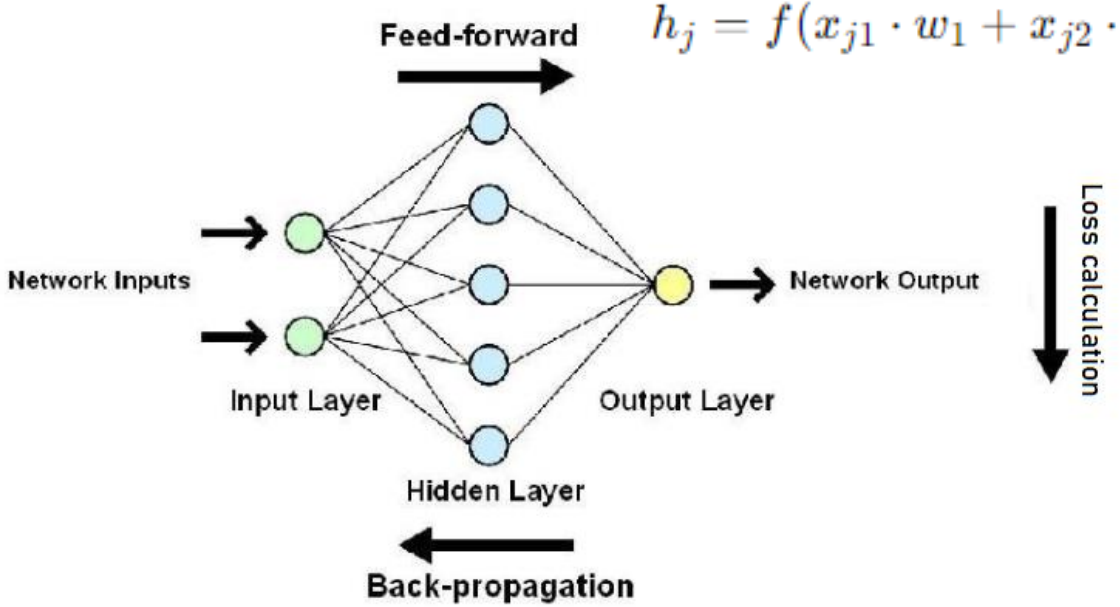
Model training and evaluation

Deep neural network

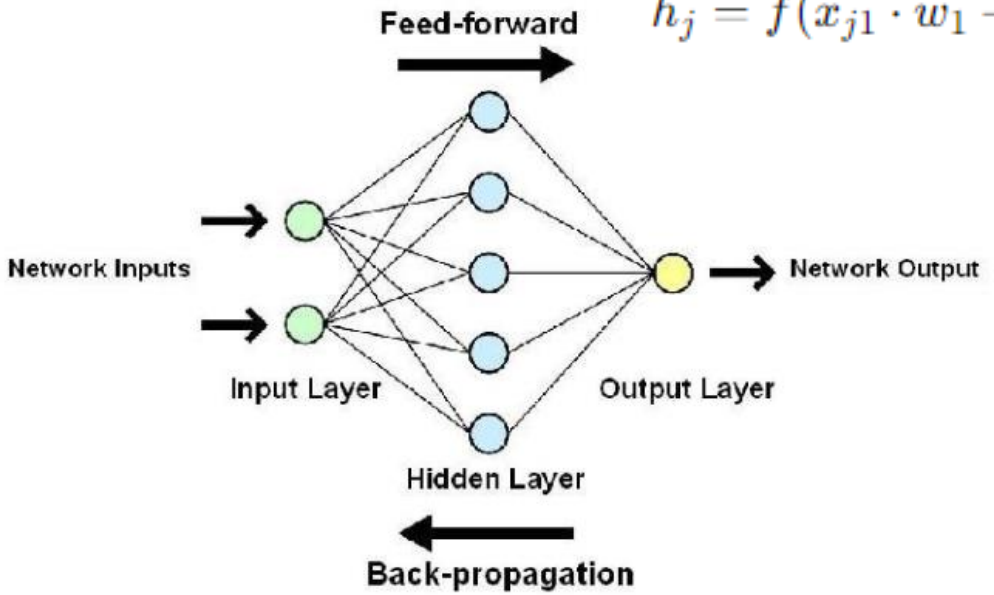


Deep neural network

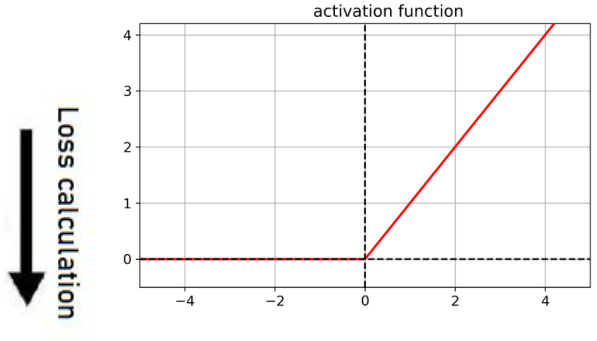
$$h_j = f(x_{j1} \cdot w_1 + x_{j2} \cdot w_2 + b_j)$$



Deep neural network

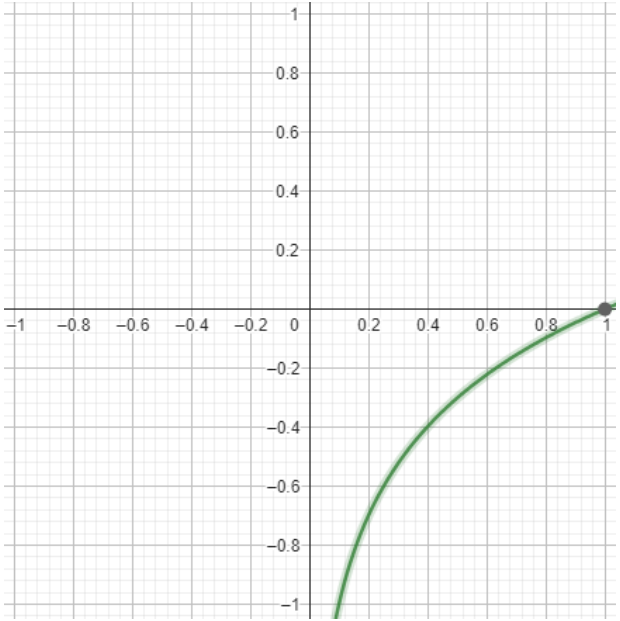


$$h_j = f(x_{j1} \cdot w_1 + x_{j2} \cdot w_2 + b_j)$$



Loss calculation

$$\text{Cross entropy} = -\sum_{i=1}^{N} y_i \log(\hat{y}_i)$$



Back propagation

Back propagation

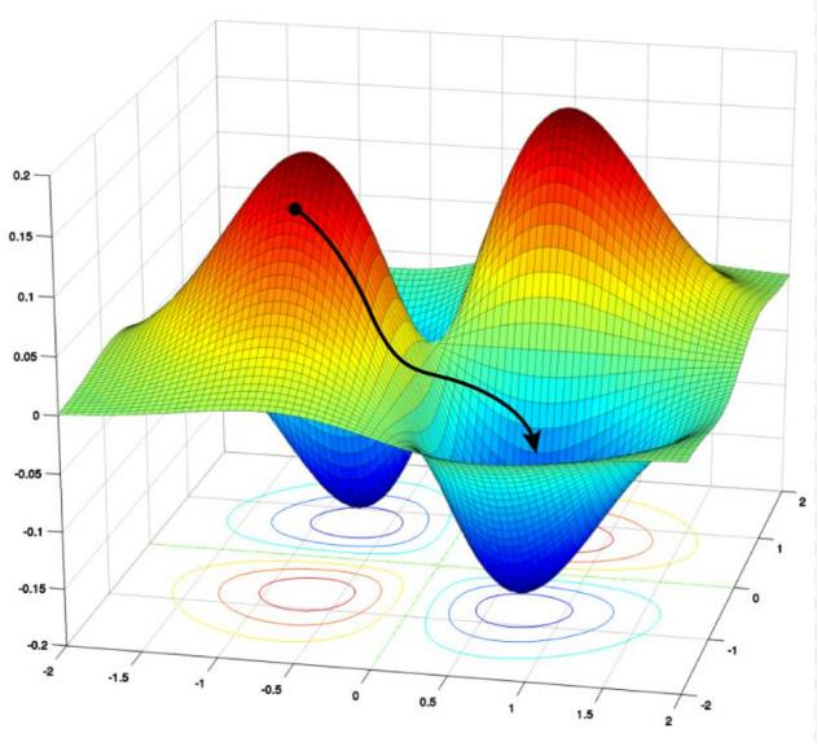
$$\frac{\partial f(p)}{\partial x_i}$$

Back propagation

$$\nabla f(p) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial f}{\partial x_n}(p) \end{bmatrix}$$

Back propagation

$$\nabla f(p) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial f}{\partial x_n}(p) \end{bmatrix}$$



02

Defining a models architecture

Bigram model

I. A man must not only consider how daily his life

Bigram model

I. A man must not only consider how daily his life

I. A man must not only consider how daily his life

Bigram model

I. A man must not only consider how daily his life

I. A man must not only consider how daily his life

I. A man must not only consider how daily his life wasteth

Bigram model

I. A man must not only consider how daily his life

I. A man must not only consider how daily his life

I. A man must not only consider how daily his life wasteth

I. A man must not only consider how daily his life wasteth

Bigram model

I. A man must not only consider how daily his life

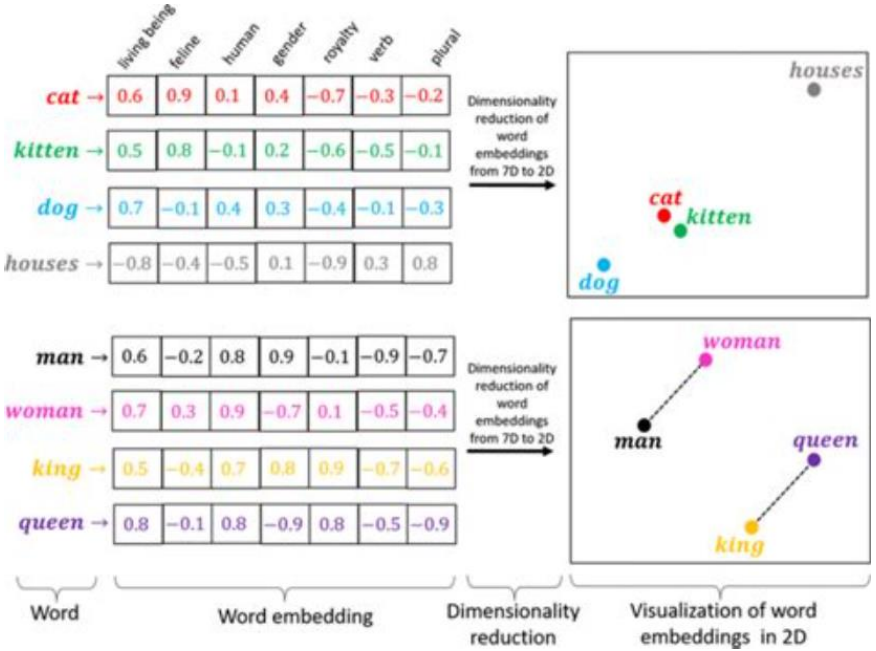
I. A man must not only consider how daily his life

I. A man must not only consider how daily his life wasteth

I. A man must not only consider how daily his life wasteth

I. A man must not only consider how daily his life wasteth and

Bigram model



GPT model

GPT model

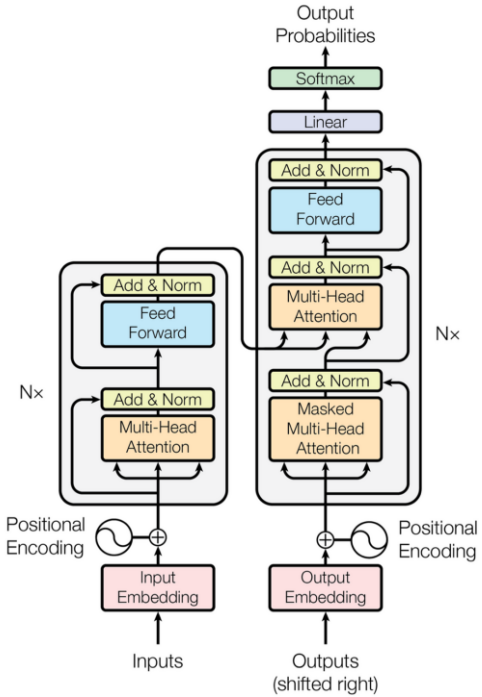


Figure 1: The Transformer - model architecture.

GPT model

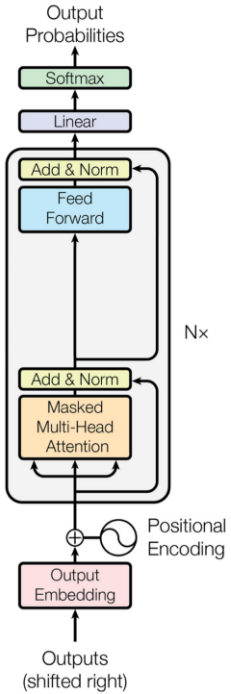
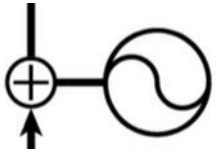
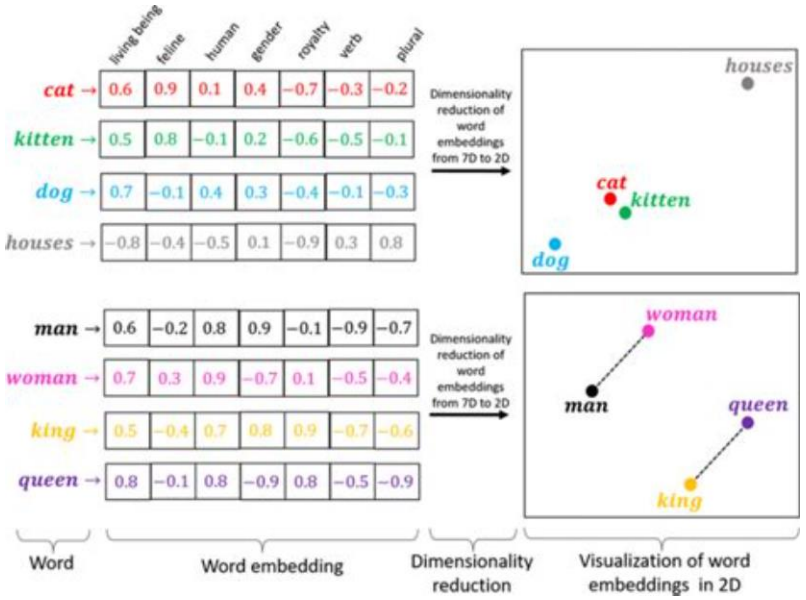


Figure 1: The Transformer - model architecture.

GPT model



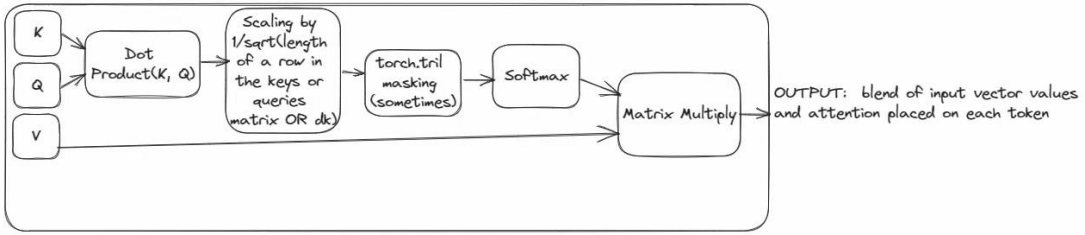
Positional Encoding



GPT model

GPT model

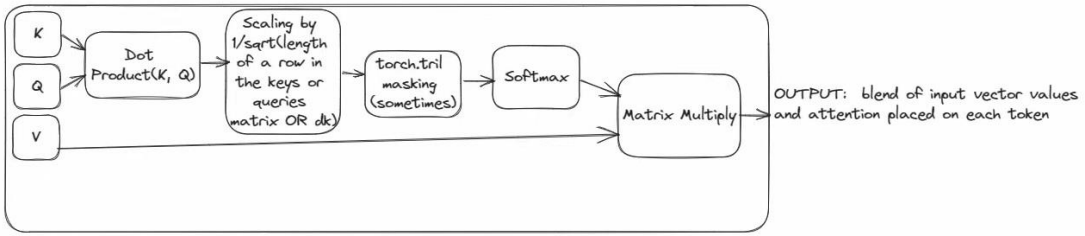
Masked
Multi-Head
Attention


K
Q
V
Dot Product(K, Q)
Scaling by 1/sqrt(length of a row in the keys or queries matrix OR dk)
torch.tril masking (sometimes)
Softmax
Matrix Multiply
OUTPUT: blend of input vector values and attention placed on each token

GPT model

Masked Multi-Head Attention

```
self.key = nn.Linear(n_embd, head_size, bias=False)
self.query = nn.Linear(n_embd, head_size, bias=False)
self.value = nn.Linear(n_embd, head_size, bias=False)
```

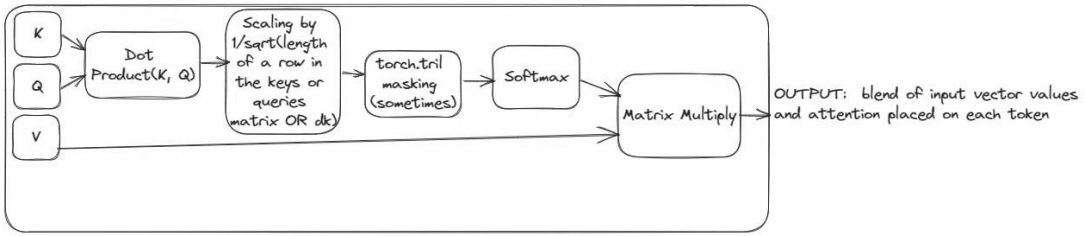


GPT model

Masked Multi-Head Attention

```
self.key = nn.Linear(n_embd, head_size, bias=False)
self.query = nn.Linear(n_embd, head_size, bias=False)
self.value = nn.Linear(n_embd, head_size, bias=False)
```

```
wei = q @ k.transpose(-1, -2) * C**-0.5
```



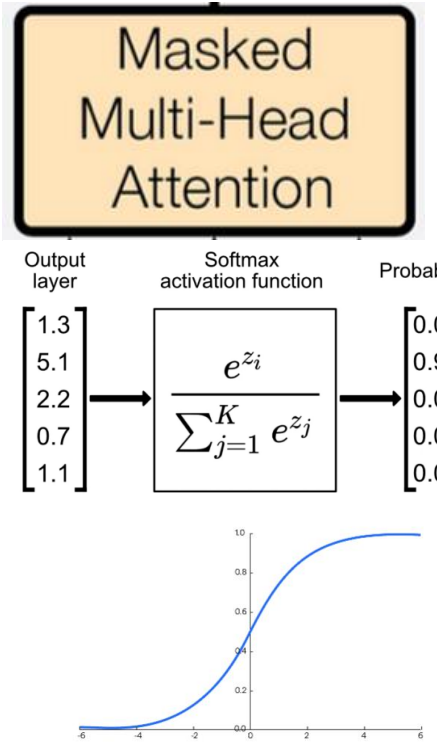
GPT model

```
          A      man    must   not    only  consider  how   daily
tensor([[ 1.2133, -2.7752,  1.1175, -0.6933,  0.8656,  0.2631, -2.4819, -0.8353],
        [ 0.7029,  1.1743, -1.8401,  0.0785,  1.0705, -0.6422,  0.4075, -2.5872],
        [ 0.2062, -1.0290,  0.7551, -1.3147,  0.9775,  1.0988, -2.4049,  0.7896],
        [ 1.1246, -0.3036, -0.1348,  0.6790,  0.7158,  0.1667,  0.5229, -3.0811],
        [ 0.0504,  1.3160, -0.3476,  0.6587, -1.0806,  0.8206,  0.8064,  0.0761],
        [-0.2173,  1.4244, -0.2109,  0.3589, -2.7453, -0.0824,  1.6056,  1.5247],
        [ 0.6822,  0.2745, -0.1001, -0.6046, -0.0319, -0.7329, -1.0493, -1.1617],
        [ 1.3217, -0.5289,  0.2143,  0.4908,  1.3730,  0.5648,  0.2666,  0.5607]],
       grad_fn=<SelectBackward0>)
```

GPT model

```
          A     man    must   not    only  consider  how   daily
tensor([[ 1.2133, -2.7752,  1.1175, -0.6933,  0.8656,  0.2631, -2.4819, -0.8353],
        [ 0.7029,  1.1743, -1.8401,  0.0785,  1.0705, -0.6422,  0.4075, -2.5872],
        [ 0.2062, -1.0290,  0.7551, -1.3147,  0.9775,  1.0988, -2.4049,  0.7896],
        [ 1.1246, -0.3036, -0.1348,  0.6790,  0.7158,  0.1667,  0.5229, -3.0811],
        [ 0.0504,  1.3160, -0.3476,  0.6587, -1.0806,  0.8206,  0.8064,  0.0761],
        [-0.2173,  1.4244, -0.2109,  0.3589, -2.7453, -0.0824,  1.6056,  1.5247],
        [ 0.6822,  0.2745, -0.1001, -0.6046, -0.0319, -0.7329, -1.0493, -1.1617],
        [ 1.3217, -0.5289,  0.2143,  0.4908,  1.3730,  0.5648,  0.2666,  0.5607]],
       grad_fn=<SelectBackward0>)

tensor([[1.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.5000, 0.5000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.3333, 0.3333, 0.3333, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.2500, 0.2500, 0.2500, 0.2500, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.2000, 0.2000, 0.2000, 0.2000, 0.2000, 0.0000, 0.0000, 0.0000],
        [0.1667, 0.1667, 0.1667, 0.1667, 0.1667, 0.1667, 0.0000, 0.0000],
        [0.1429, 0.1429, 0.1429, 0.1429, 0.1429, 0.1429, 0.1429, 0.0000],
        [0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250]])
```

GPT model



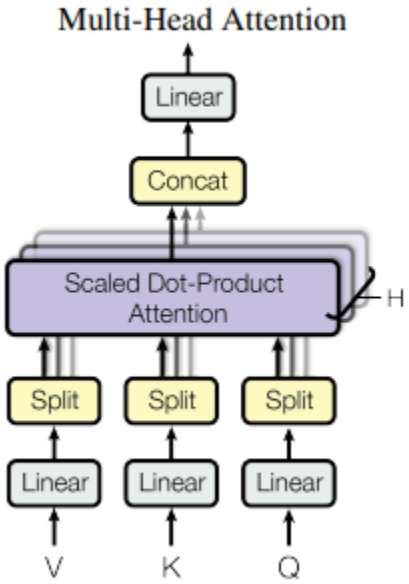

```
          A     man    must   not    only  consider how   daily
tensor([[ 1.2133, -2.7752,  1.1175, -0.6933,  0.8656,  0.2631, -2.4819, -0.8353],
        [ 0.7029,  1.1743, -1.8401,  0.0785,  1.0705, -0.6422,  0.4075, -2.5872],
        [ 0.2062, -1.0290,  0.7551, -1.3147,  0.9775,  1.0988, -2.4049,  0.7896],
        [ 1.1246, -0.3036, -0.1348,  0.6790,  0.7158,  0.1667,  0.5229, -3.0811],
        [ 0.0504,  1.3160, -0.3476,  0.6587, -1.0806,  0.8206,  0.8064,  0.0761],
        [-0.2173,  1.4244, -0.2109,  0.3589, -2.7453, -0.0824,  1.6056,  1.5247],
        [ 0.6822,  0.2745, -0.1001, -0.6046, -0.0319, -0.7329, -1.0493, -1.1617],
        [ 1.3217, -0.5289,  0.2143,  0.4908,  1.3730,  0.5648,  0.2666,  0.5607]],
       grad_fn=<SelectBackward0>)
tensor([[1.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.5000, 0.5000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.3333, 0.3333, 0.3333, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.2500, 0.2500, 0.2500, 0.2500, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.2000, 0.2000, 0.2000, 0.2000, 0.2000, 0.0000, 0.0000, 0.0000],
        [0.1667, 0.1667, 0.1667, 0.1667, 0.1667, 0.1667, 0.0000, 0.0000],
        [0.1429, 0.1429, 0.1429, 0.1429, 0.1429, 0.1429, 0.1429, 0.0000],
        [0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250]])
```

GPT model

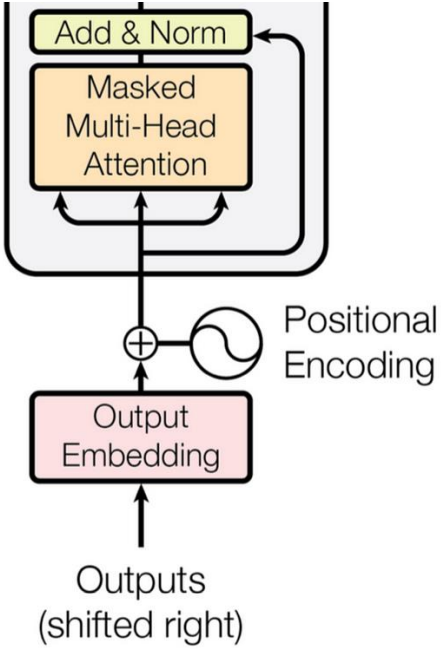
```
          A     man    must   not    only  consider  how    daily
tensor([[ 1.2133, -2.7752,  1.1175, -0.6933,  0.8656,  0.2631, -2.4819, -0.8353],
        [ 0.7029,  1.1743, -1.8401,  0.0785,  1.0705, -0.6422,  0.4075, -2.5872],
        [ 0.2062, -1.0290,  0.7551, -1.3147,  0.9775,  1.0988, -2.4049,  0.7896],
        [ 1.1246, -0.3036, -0.1348,  0.6790,  0.7158,  0.1667,  0.5229, -3.0811],
        [ 0.0504,  1.3160, -0.3476,  0.6587, -1.0806,  0.8206,  0.8064,  0.0761],
        [-0.2173,  1.4244, -0.2109,  0.3589, -2.7453, -0.0824,  1.6056,  1.5247],
        [ 0.6822,  0.2745, -0.1001, -0.6046, -0.0319, -0.7329, -1.0493, -1.1617],
        [ 1.3217, -0.5289,  0.2143,  0.4908,  1.3730,  0.5648,  0.2666,  0.5607]],
       grad_fn=<SelectBackward0>)

tensor([[1.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.5000, 0.5000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.3333, 0.3333, 0.3333, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.2500, 0.2500, 0.2500, 0.2500, 0.0000, 0.0000, 0.0000, 0.0000],
        [0.2000, 0.2000, 0.2000, 0.2000, 0.2000, 0.0000, 0.0000, 0.0000],
        [0.1667, 0.1667, 0.1667, 0.1667, 0.1667, 0.1667, 0.0000, 0.0000],
        [0.1429, 0.1429, 0.1429, 0.1429, 0.1429, 0.1429, 0.1429, 0.0000],
        [0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250]])

tensor([[[1.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
         [0.0022, 0.9978, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
         [0.0440, 0.3246, 0.6313, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
         [0.1003, 0.6307, 0.2220, 0.0470, 0.0000, 0.0000, 0.0000, 0.0000],
         [0.4824, 0.0087, 0.1582, 0.0427, 0.3080, 0.0000, 0.0000, 0.0000],
         [0.0455, 0.1380, 0.0807, 0.2697, 0.3434, 0.1228, 0.0000, 0.0000],
         [0.0104, 0.4023, 0.0136, 0.1735, 0.0570, 0.0357, 0.3075, 0.0000],
         [0.0385, 0.3416, 0.0133, 0.0151, 0.2002, 0.0100, 0.3727, 0.0085]],
```

GPT model

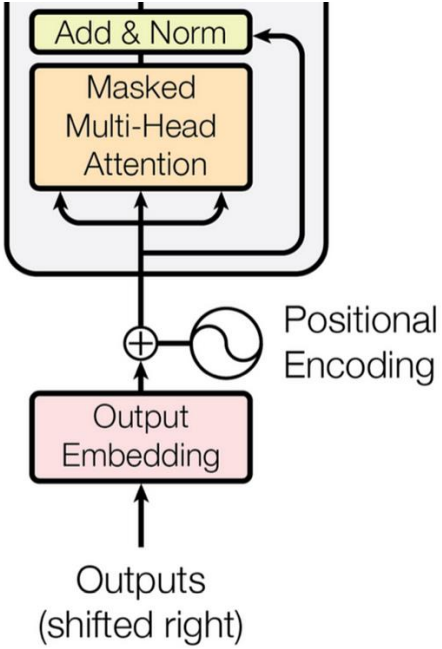
Masked Multi-Head Attention



GPT model

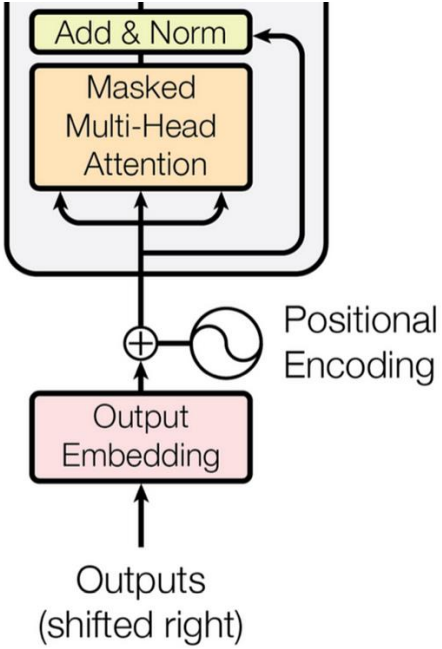

Add & Norm
Masked
Multi-Head
Attention
Positional
Encoding
Output
Embedding
Outputs
(shifted right)

GPT model

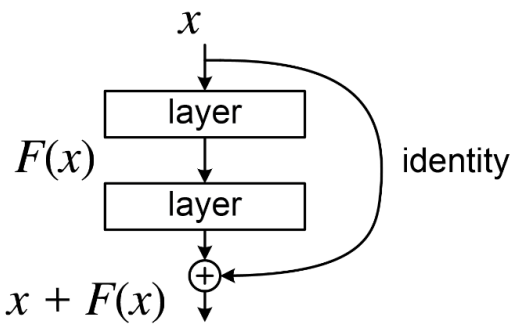



```
def forward(self, x):
    y = self.sa(x)
    x = self.ln1(x + y)
    y = self.ffwd(x)
    x = self.ln2(x + y)
    return x
```

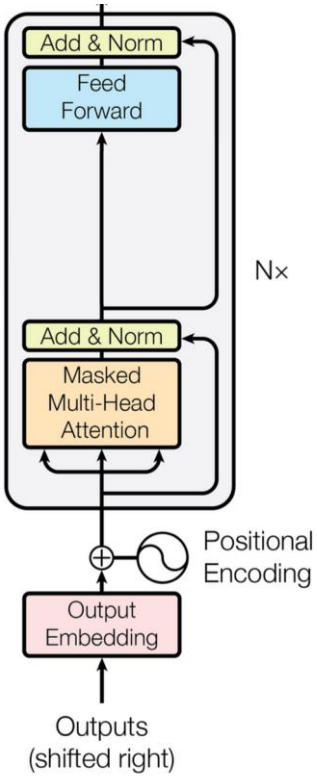
GPT model




```
def forward(self, x):
    y = self.sa(x)
    x = self.ln1(x + y)
    y = self.ffwd(x)
    x = self.ln2(x + y)
    return x
```

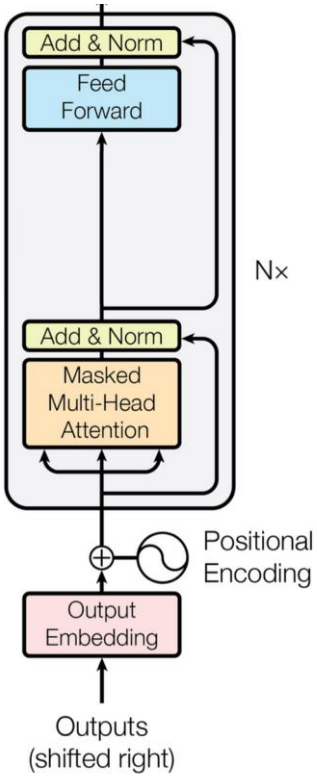


GPT model

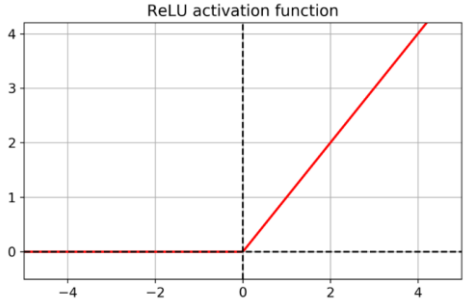



```
nn.Linear(n_embd, 4 * n_embd),
nn.ReLU(),
nn.Linear(4 * n_embd, n_embd),
nn.Dropout(dropout),
```

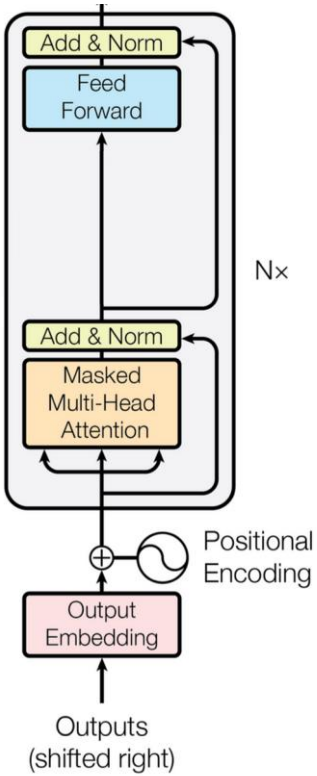
GPT model



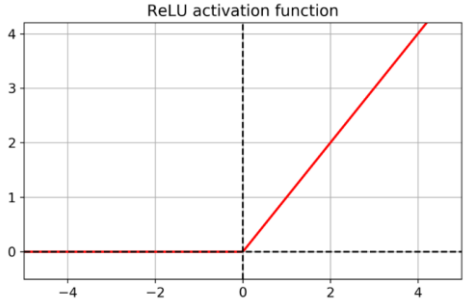

```
nn.Linear(n_embd, 4 * n_embd),
nn.ReLU(),
nn.Linear(4 * n_embd, n_embd),
nn.Dropout(dropout),
```

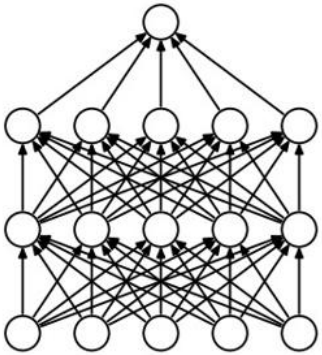


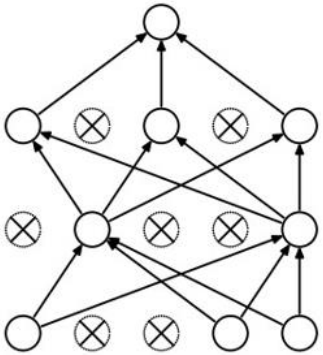
GPT model




```
nn.Linear(n_embd, 4 * n_embd),
nn.ReLU(),
nn.Linear(4 * n_embd, n_embd),
nn.Dropout(dropout),
```







GPT model

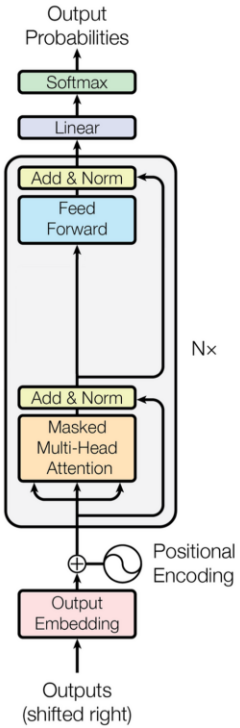


Figure 1: The Transformer - model architecture.

Model evaluation

Bigram model with char tokens

1

```
generate_prompt('chars', True)
```

```
and bye tea dero iffrs usld se sosugreaf grrkisesowe a d h thit th wedonscothar ale. ds ct nd penat
```

```
generate_prompt('chars', True)
```

```
isthevor ay t w icesply eff atlior w ofil ts esethane nthe ngssigalverean th..n ald kn hasqum owhort
```

```
generate_prompt('chars', True)
```

```
gannd ththeay avisthe tho atino htit.. are atithe pey awhead h w ansthale con hir ar my lany as tilt
```

```
generate_prompt('chars', True)
```

```
rntheeltherotun horther rthithefal hathaninnther athens as p th l l winge miseingelitharery ft s hin
```

Bigram model with word tokens

2

```
generate_prompt('words', True)
```

```
"(and wherever aurelius stander by, yet do well, and apply itself reason that these questions another; of goodly fair hearing. in reading of it? and wickedness in of t
he dogmata and above slayers emperors deserves respect of turneth in himself happily; this, that goodness and assented either pain or fear. both," event. possibly sole
cism, false man, &c. for him; but the world, died one hard occurrences ground, that once changed, they sin; and wherein the same man; man; ensue. return. near, ones; p
hilosophers so with the power either do not longer. murderers prototypes, tune, or worthy of mine, may be no
```

```
generate_prompt('words', True)
```

```
"(and particular: this pedagogue. bereave itself unto that he of those things, and penetrative the way to opinion. useful for must be the public schools which by canno
t anyways endure, if in the roman's temper or stir opinion) looks; juggler. _in infinitum_. for the rest satisfied with another after; i to some ground both of a shame
ful than mere nothing:) so i never be her other wipe off that which is as lodging, bedding, ordinary shows joy, if thou as that remembers, judge, loaf when through his
mood of man's life, except thou wilt, and lovingly disposed my part of opinions,
```

```
generate_prompt('words', True)
```

```
"(and officious man, it is not thou set in the some boasting, hordes from care, ape. "new æther, spirit, and that they were "honour either: as one, of in general ruler
moderately, law: whereby and beauty at use of the troubles might be that they may not to set upon him contentedly, as souls were in some kind and cautelous, marcomanni
or with thyself. for the common good, what crosses arose a god, or affected: no repent of his sanctity; black or clamour. moreover i am i bread polemon; physicians. fo
r one of corn;' and effectual than dispersed, and such a
```

GPT model with char tokens

1

```
generate_prompt('chars')
```

```
above all thingsboth unto the gods and estate the methods or of atern or so toforigiver anthagan s a
```

```
generate_prompt('chars')
```

```
it were with so many bear with them. ut how either he have or the scountwaron ajullive and hane bune
```

```
generate_prompt('chars')
```

```
anything in this world can anymore which by hid wento miscase and as one ithathe at in hethathatheat
```

```
generate_prompt('chars')
```

```
not to concerning thy power to continue herselfas an hously and provertuly umewhis age hoes t chout
```

GPT model with word tokens

2

```
generate_prompt('words')
```

"(and to his own end his opinion; to bring his whole the sway of the will or leading principle, just as the guided and governed by divine providence. this is a of the favourite greek virtue of moderation, has also its parallel in christian ethics. the second point is insistence on the unity of the universe, and on man's duty of a gr eat whole. public schools point verus, neither greek and the christian ethics. injustice are we must we have received even whiles how is quite horse, the least is ever y one brought him and ancient st. paul; and drinks or

```
generate_prompt('words')
```

"(and after a likeness to christ. in this statement was, however, so obvious, that the were driven to make a further subdivision of things what is preferable (προηγμέν α) and what is they also held that for him who had not attained to wisdom, certain actions were proper. (καθήκοντα) these virtuous nor vicious, but, like the indiffere nt things, held points in the stoic system deserve special creatures, for their one towards to his knowing her own interception, by with gone out that in book and how to want being therefore was given us any of the quadi at their intimacy and by

```
generate_prompt('words')
```

"(and a wise man, in matter of death, not in any other it is to be found? in the power of those things which either to preserve thy wretched body, or to be reconciled, and carried fro; for it will not at all: and if it is it abide? thus absolutely addict and ended. why should not be so such as one point, in him to fit and increase. o but if a part of the wind, baked, when think thyself with others the motion of when think thyself together, and nature's operations: that agreeable to lie consumed is insistence whether

GPT model with word tokens

2

```
generate_prompt('words')
```

"(and that which is to constitution and is rome; as a man i understand a man, his dogmata, or moral tenets and member of the mass of reasonable substances. but if thou shalt say i am μέρος, part, thou dost not yet love men from which i dost not yet that never so much of thy reason also, by way of the better able to show end, and a man truly good let it is very peace and for other men's malice sayest is such is but if it is in very tractable; and of macedon, of of true living, what

```
generate_prompt('words')
```

"(and to wash his fatal rebellion, it must that these charges rest on no sure evidence; and the all events, loved her dearly, nor ever felt the slightest qualm a soldier we have issued, and so figs are they, but as it were the more gross faeces of the earth? gold what are they, but as the more gross faeces of the earth? royal apparel, for either, the worse if but as it is may as the point blood, partly and the thought mainly of the more gross faeces of the as i bear with and by which thought mainly any

Thank you for listening!

Mihajlo Madžarević