

gpt

January 9, 2026

1 Construction d'un GPT

1.1 Introduction

Nous allons dans ce cours construire “à la main” un réseau de neurones correspondant à un **GPT** (generative pre-trained transformer).

- [V+2017] Vaswany, A. *et al.* “[Attention Is All You Need](#)”, 2017 (Transformer)
- [R+2018] Radford, A. *et al.* “[Improving Language Understanding by Generative Pre-Training](#)”, 2018 (GPT-1)
- [R+2019] Radford, A. *et al.* “[Language Models are Unsupervised Multitask Learners](#)”, 2019 (GPT-2)
- [B+2020] Brown T. B. *et al.* “[Language Models are Few-Shot Learners](#)”, 2020, (GPT-3)
- [O+2020] [OpenAI ChatGPT blog post](#)
- Une vidéo très bien faite de [3Blue1Brown](#) sur le mécanisme d’attention: [Attention in transformers, visually explained](#)
- La vidéo d’Andrej Karpathy “[Let’s build GPT from scratch](#)” dont est issu ce cours
 - [nanoGPT](#)
 - [llm.c](#)
- [LoshchilovI2017] Loshchilov, I. et Hutter, F. “[Decoupled Weight Decay Regularization](#)”, 2017 (AdamW)

1.2 Données

1.2.1 Imports

Contrairement à ce que nous avions fait jusqu’à présent, nous allons maintenant étendre des modules de PyTorch plutôt que de définir nos propres classes.

Nous définissons également ici un paramètre `device` qui va nous permettre de faire tourner le code sur:

- le CPU, c’est-à-dire le processeur de votre machine (`cpu`)
- `cuda` pour un GPU NVIDIA
- `mps` pour un mac sous macOS avec un processeur Arm de type M1, M2, M3 ou M4

```
[1]: import torch
import torch.nn as nn
from torch.nn import functional as F
device = 'cpu'
# device = 'cuda'
# device = 'mps'
seed = 2147483647
torch.manual_seed(seed);
```

1.2.2 Données texte d'entrée

Reprenons ici la classe `TextData` que nous avions utilisée dans les précédents Notebooks. Nous la simplifions pour y intégrer directement les données d'entraînement et de validation.

Le constructeur de la classe a également un nouveau paramètre `device` qui permet de préciser où sera stocké le tenseur contenant les données.

```
[2]: class TextData(object):

    def __init__(self, filename, device):
        self.filename = filename
        self.device = device
        with open(self.filename, 'r', encoding="utf-8") as f:
            self.text = f.read()
        self.size = len(self.text)
        self.chars = sorted(list(set(self.text)))
        self.vocab_size = len(self.chars)
        self.ctoi = {c:i for i,c in enumerate(self.chars)}
        self.itoc = {i:c for c,i in self.ctoi.items()}
        self.encode = lambda s: [self.ctoi[c] for c in s]
        self.data = torch.tensor(self.encode(self.text), dtype=torch.long, ↴
        device=device)
        n = int(0.9*len(self.data))
        self.train_data = self.data[:n]
        self.val_data = self.data[n:]

    def decode(self, block):
        if isinstance(block, torch.Tensor):
            l = block.tolist()
            # !\ https://docs.pytorch.org/docs/stable/generated/torch.Tensor.
            ↴tolist.html
            if isinstance(l, int):
                l = [l]
            elif isinstance(block, int):
                l = [block]
            else:
                l = block
            assert isinstance(l, list)
```

```

    return ''.join([self.itoc[i] for i in l])

def _get_batch(self, data, batch_size, block_size, device):
    if device is None:
        device = self.device
    ix = torch.randint(len(data) - block_size, (batch_size,))
    x = torch.stack([data[i : i + block_size] for i in ix])
    y = torch.stack([data[i + 1 : i + block_size + 1] for i in ix])
    x, y = x.to(device), y.to(device)
    return x, y

def get_training_batch(self, batch_size, block_size, device=None):
    data = self.train_data
    return self._get_batch(data, batch_size, block_size, device)

def get_validation_batch(self, batch_size, block_size, device=None):
    data = self.val_data
    return self._get_batch(data, batch_size, block_size, device)

def __repr__(self):
    l = []
    chars_str = ''.join(self.chars)
    l.append("<TextData")
    l.append(f' filename=\"{self.filename}\"")
    l.append(f' device=\"{self.device}\"")
    l.append(f' size=\"{self.size}\"")
    l.append(f' vocab_size=\"{self.vocab_size}\">")
    return '\n'.join(l)

```

1.2.3 Exemple du code civil français

Dans ce notebook, nous allons utiliser les données du [code civil français](#) provenant du site Légifrance, transformé en markdown.

Si vous n'avez pas récupéré le fichier `civil.md` depuis la page du cours, vous pouvez le récupérer localement en décommentant les deux lignes suivantes. Il faut que la commande `wget` soit installée sur votre machine.

```
[ ]: # !wget https://storage.gra.cloud.ovh.net/v1/
  ↪AUTH_4d7d1bcd41914ee184ef80e2c75c4fb1/dila-legi-codes/civil.md.zip
# !python -m zipfile -e civil.md.zip civil_test.md
```

```
[3]: with open('civil.md', 'r', encoding="utf-8") as f:
    code_civil = f.read()
    print(code_civil[:1500])
```

```
---
title: Code civil
date: 2024-01-15
```

```
## Titre préliminaire : De la publication, des effets et de l'application des  
lois en général
```

****Art. 1****

Les lois et, lorsqu'ils sont publiés au Journal officiel de la République française, les actes administratifs entrent en vigueur à la date qu'ils fixent ou, à défaut, le lendemain de leur publication. Toutefois, l'entrée en vigueur de celles de leurs dispositions dont l'exécution nécessite des mesures d'application est reportée à la date d'entrée en vigueur de ces mesures.

En cas d'urgence, entrent en vigueur dès leur publication les lois dont le décret de promulgation le prescrit et les actes administratifs pour lesquels le Gouvernement l'ordonne par une disposition spéciale.

Les dispositions du présent article ne sont pas applicables aux actes individuels.

****Art. 2****

La loi ne dispose que pour l'avenir ; elle n'a point d'effet rétroactif.

****Art. 3****

Les lois de police et de sûreté obligent tous ceux qui habitent le territoire.

Les immeubles, même ceux possédés par des étrangers, sont régis par la loi française.

Les lois concernant l'état et la capacité des personnes régissent les Français, même résidant en pays étranger.

****Art. 4****

Le juge qui refusera de juger, sous prétexte du silence, de l'obscurité ou de l'insuffisance de la loi, pourra être poursuivi comme coupable de déni de justice.

****Art. 5****

Il est défendu aux juges de prononcer par voie de disposition générale et réglementaire sur les causes qui leur son

```
[4]: text = TextData('civil.md', device)  
print(text)
```

```
<TextData
  filename="civil.md"
  device="cpu"
  size="1182082"
  vocab_size="91"/>
```

Comme précédemment, nous allons travailler sur un modèle de langue par caractères. Notre vocabulaire va donc être composé des `vocab_size` caractères du code civil (91).

[5] : `print(text.chars)`

```
['\n', ' ', "'", '#', '%', "", '(', ')', '*', ',', '-', '.', '0', '1', '2',
'3', '4', '5', '6', '7', '8', '9', ':', ';', 'A', 'B', 'C', 'D', 'E', 'F', 'G',
'H', 'I', 'J', 'L', 'M', 'N', 'O', 'P', 'Q', 'R', 'S', 'T', 'U', 'V', 'W', 'X',
'Y', 'a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j', 'l', 'm', 'n', 'o', 'p',
'q', 'r', 's', 't', 'u', 'v', 'x', 'y', 'z', '\xa0', '°', 'É', 'à', 'â', 'ç',
'è', 'é', 'ê', 'ë', 'î', 'ï', 'ô', 'ù', 'œ', '‐', '‐', '€']
```

Les méthodes `encode` et `decode` vont nous permettre de transformer chaque caractère en son numéro de token (un entier) et de transformer chaque token en son caractère.

[7] : `# Article 2 du code civil`

```
print(text.encode("La loi ne dispose que pour l'avenir ; elle n'a point d'effet
  ↪rétroactif.\n"))
```

```
[34, 48, 1, 58, 61, 56, 1, 60, 52, 1, 51, 56, 65, 62, 61, 65, 52, 1, 63, 67, 52,
1, 62, 61, 67, 64, 1, 58, 5, 48, 68, 52, 60, 56, 64, 1, 23, 1, 52, 58, 58, 52,
1, 60, 5, 48, 1, 62, 61, 56, 60, 66, 1, 51, 5, 52, 53, 53, 52, 66, 1, 64, 79,
66, 64, 61, 48, 50, 66, 56, 53, 11, 0]
```

[8] : `print(text.decode(text.encode("La loi ne dispose que pour l'avenir ; elle n'a`

```
  ↪point d'effet rétroactif.\n")))
```

La loi ne dispose que pour l'avenir ; elle n'a point d'effet rétroactif.

Nous ne le ferons pas ici, mais nous aurions également pu utiliser un tokenizer existant, comme celui de GPT2, mais dans le cas notre vocabulaire serait beaucoup plus grand.

[12] : `# !pip install tiktoken`

```
import tiktoken
enc = tiktoken.get_encoding('gpt2')
print(enc.n_vocab)
print(enc.encode("La loi ne dispose que pour l'avenir ; elle n'a point d'effet
  ↪rétroactif."))
print(enc.decode([14772]))
print(enc.decode([14772, 2376]))
print(enc.decode([14772, 2376, 72]))
print(enc.decode([14772, 2376, 72, 497]))
print(enc.decode([14772, 2376, 72, 497, 34291]))
```

```
#print(enc.decode([14772, 2376, 72, 497, 34291, 8358, 12797, 300, 6, 4005, 343, 2162, 1288, 293, 299, 6, 64, 966, 288, 6, 14822, 316, 40560, 23528, 529, 361, 13]))
```

```
50257  
[14772, 2376, 72, 497, 34291, 8358, 12797, 300, 6, 4005, 343, 2162, 1288, 293,  
299, 6, 64, 966, 288, 6, 14822, 316, 40560, 23528, 529, 361, 13]  
La  
La lo  
La loi  
La loi ne  
La loi ne dispose
```

Les données que nous allons utiliser contiennent près de 1,2 millions de tokens.

```
[9]: print(text.data.shape)  
      print(text.data[:100])
```

```
torch.Size([1182082])  
tensor([10, 10, 10, 0, 66, 56, 66, 58, 52, 22, 1, 26, 61, 51, 52, 1, 50, 56,  
       68, 56, 58, 0, 51, 48, 66, 52, 22, 1, 14, 12, 14, 16, 10, 12, 13, 10,  
       13, 17, 0, 10, 10, 10, 0, 0, 3, 3, 1, 42, 56, 66, 64, 52, 1, 62,  
       64, 79, 58, 56, 59, 56, 60, 48, 56, 64, 52, 1, 22, 1, 27, 52, 1, 58,  
       48, 1, 62, 67, 49, 58, 56, 50, 48, 66, 56, 61, 60, 9, 1, 51, 52, 65,  
       1, 52, 53, 53, 52, 66, 65, 1, 52, 66])
```

1.3 Données d'entraînement pour modèle causal: batch, contexte et embeddings (B, T, C)

1.3.1 Time: T (block_size)

Comme précédemment, nous n'allons pas entraîner le modèle sur toutes les données, mais seulement sur des portions de données contigües de taille fixe à la fois, appelées *contexte* ou *bloc*. Dans le reste de ce document, nous allons définir la taille de ce morceau avec un paramètre `block_size`.

```
[10]: block_size = 8
```

Chacun de ces blocs de données va nous permettre de générer des exemples d'entraînement: dans notre cadre de modèles génératifs, un certain nombre de tokens et le token suivant. À la différence de ce que nous avions fait jusqu'à présent, nous allons avec cette architecture de type Transformer faire "communiquer" ces tokens entre-eux d'une manière différente avec le mécanisme d'attention (que nous verrons plus loin). Ainsi, avec un contexte d'une certaine taille maximale, nous allons exploiter tous les tokens de ce contexte pour générer des exemples.

Par exemple, si l'on prend un bloc de `block_size` tokens à partir de l'offset 151 comme données source `x`:

```
[11]: offset = 151  
x = text.train_data[offset : offset + block_size]  
print(x)
```

```
text.decode(x)

tensor([34, 52, 65, 1, 58, 61, 56, 65])
```

[11]: 'Les lois'

et un bloc de 8 tokens décalés de 1 token:

```
[13]: y = text.train_data[offset + 1 : offset + block_size + 1]
print(y)
text.decode(y)

tensor([52, 65, 1, 58, 61, 56, 65, 1])
```

[13]: 'es lois '

On obtient ainsi 8 (`block_size`) exemples sur lesquels on va pouvoir entraîner notre modèle:

```
[14]: for t in range(block_size):
    context = x[:t+1]
    target = y[t]
    print(f"t_{t} : '{text.decode(context)}' -> '{text.decode(target)}'"
          f"\n{context.tolist()} -> {target}")

t_0 : 'L' -> 'e' ([34] -> 52)
t_1 : 'Le' -> 's' ([34, 52] -> 65)
t_2 : 'Les' -> '' ([34, 52, 65] -> 1)
t_3 : 'Les ' -> 'l' ([34, 52, 65, 1] -> 58)
t_4 : 'Les 1' -> 'o' ([34, 52, 65, 1, 58] -> 61)
t_5 : 'Les lo' -> 'i' ([34, 52, 65, 1, 58, 61] -> 56)
t_6 : 'Les loi' -> 's' ([34, 52, 65, 1, 58, 61, 56] -> 65)
t_7 : 'Les lois' -> '' ([34, 52, 65, 1, 58, 61, 56, 65] -> 1)
```

Le `t` dans la boucle précédente représente un *pas de temps*, lorsque l'on augmente progressivement la taille du contexte, en passant de t_0 à t_7 . Dans la suite de ce notebook, nous représenterons de manière courte la taille de cette fenêtre de temps, soit `block_size` par la lettre T (comme *time*). `block_size` représente donc la taille maximale du contexte que nous utiliserons pour les prédictions.

1.3.2 Batch: B (`batch_size`)

Lors de l'entraînement de notre modèle, nous allons utiliser à chaque étape plusieurs exemples indépendants, ensemble appelé un *batch* de taille `batch_size`, que nous abrégerons dans la suite par la lettre B. Ces exemples seront comme précédemment tirés aléatoirement avec `torch.randint` dans notre jeu de données d'entraînement. Chaque exemple sera de taille T (`block_size`).

```
[15]: batch_size = 4
ix = torch.randint(len(text.data) - block_size, (batch_size,))
print(ix)

tensor([ 418285, 46560, 1094659, 1152818])
```

Nous allons utiliser ces indices pour empiler (“stacker”) B tenseurs de taille T avec `torch.stack` sur la dimension 0 pour nos x et y:

```
[16]: x = torch.stack([text.data[i : i + block_size] for i in ix])
y = torch.stack([text.data[i + 1 : i + block_size + 1] for i in ix])
print(x.shape) # B x T
print(y.shape) # B x T
```

```
torch.Size([4, 8])
torch.Size([4, 8])
```

C'est ce principe qui est utilisé dans la classe `TextData` ci-dessus dans les méthodes `get_training_data` et `get_validation_data`, dont voici une illustration:

```
[17]: xb, yb = text.get_training_batch(batch_size, block_size)
print('inputs (B, T):')
print(xb.shape)
print(xb)
print('targets (B, T):')
print(yb.shape)
print(yb)

print('----')

for b in range(batch_size): # batch dimension
    for t in range(block_size): # time dimension
        context = xb[b, :t+1]
        target = yb[b,t]
        print(f"when input is {context.tolist()} the target: {target}")
```

```
inputs (B, T):
torch.Size([4, 8])
tensor([[65, 56, 61, 60, 1, 51, 52, 1],
       [51, 52, 65, 1, 62, 64, 52, 65],
       [59, 52, 59, 49, 64, 52, 65, 9],
       [52, 65, 1, 79, 62, 61, 67, 69]])
targets (B, T):
torch.Size([4, 8])
tensor([[56, 61, 60, 1, 51, 52, 1, 58],
       [52, 65, 1, 62, 64, 52, 65, 66],
       [52, 59, 49, 64, 52, 65, 9, 1],
       [65, 1, 79, 62, 61, 67, 69, 9]])
-----
when input is [65] the target: 56
when input is [65, 56] the target: 61
when input is [65, 56, 61] the target: 60
when input is [65, 56, 61, 60] the target: 1
when input is [65, 56, 61, 60, 1] the target: 51
when input is [65, 56, 61, 60, 1, 51] the target: 52
```

```

when input is [65, 56, 61, 60, 1, 51, 52] the target: 1
when input is [65, 56, 61, 60, 1, 51, 52, 1] the target: 58
when input is [51] the target: 52
when input is [51, 52] the target: 65
when input is [51, 52, 65] the target: 1
when input is [51, 52, 65, 1] the target: 62
when input is [51, 52, 65, 1, 62] the target: 64
when input is [51, 52, 65, 1, 62, 64] the target: 52
when input is [51, 52, 65, 1, 62, 64, 52] the target: 65
when input is [51, 52, 65, 1, 62, 64, 52, 65] the target: 66
when input is [59] the target: 52
when input is [59, 52] the target: 59
when input is [59, 52, 59] the target: 49
when input is [59, 52, 59, 49] the target: 64
when input is [59, 52, 59, 49, 64] the target: 52
when input is [59, 52, 59, 49, 64, 52] the target: 65
when input is [59, 52, 59, 49, 64, 52, 65] the target: 9
when input is [59, 52, 59, 49, 64, 52, 65, 9] the target: 1
when input is [52] the target: 65
when input is [52, 65] the target: 1
when input is [52, 65, 1] the target: 79
when input is [52, 65, 1, 79] the target: 62
when input is [52, 65, 1, 79, 62] the target: 61
when input is [52, 65, 1, 79, 62, 61] the target: 67
when input is [52, 65, 1, 79, 62, 61, 67] the target: 69
when input is [52, 65, 1, 79, 62, 61, 67, 69] the target: 9

```

Remarque: lors de l'inférence, B sera égal à 1 puisqu'on a un contexte unique.

1.3.3 Channels: C (n_embd)

Chaque token parmi les T du contexte sera représenté par un plongement lexical dans un espace à n_embd dimensions, que nous appelerons “*channels*” et qui sera matérialisé par la lettre C.

1.4 Retour au bigramme

1.4.1 NeuronalBigram(nn.Module)

Nous allons dans cette section reprendre le bigramme neuronal que nous avions construit lors des première séance. Cela va nous permettre d'illustrer l'utilisation de la classe `torch.nn.module` qui est la classe de base pour tous les modules composants les réseaux de neurones implémentés avec PyTorch. Parmi les aspects intéressants de cette classe, il y a le fait de pouvoir définir d'autres modules en attribut ou encore de pouvoir redéfinir la méthode `forward` qui sera appelée lors de la passe de propagation avant.

Définissons une classe `NeuronalBigram` qui va simplement contenir un dictionnaire `token_embedding_table`, de type `torch.nn.Embedding` de `vocab_size` entrées, où chaque entrée sera un tenseur à une dimension de taille `vocab_size`, similaire à ce que nous avions défini dans le troisième cours:

```
W = torch.randn((nb_chars, nb_chars), generator=g, requires_grad=True)
```

```
[18]: class NeuronalBigram(nn.Module):
```

```
    def __init__(self, vocab_size, device):
        super().__init__()
        self.device = device
        # Dictionnaire de `vocab_size` entrées où chaque entrée est un tenseur
        # de dimension 1 de taille `vocab_size`
        self.token_embedding_table = nn.Embedding(vocab_size, vocab_size,
                                                device=device)
```

1.4.2 Passe avant

Nous pouvons maintenant définir la méthode `forward` qui va permettre de définir le calcul effectué lorsqu'on appelle le modèle. Comme nous aurons également besoin d'un calcul du "loss" lors de l'évaluation et également d'effectuer le calcul lors de l'inférence (génération de mots), cette méthode va prendre un paramètre optionnel `target` permettant de définir la solution de référence pour l'évaluation du loss. Nous devons également réorganiser les `logits` car la fonction `cross_entropy` attend une entrée de la forme (B, C) pour le premier paramètres et (B) pour le deuxième.

```
[19]: def forward(self, idx, targets=None):
```

"""Calcul effectué par le modèle.

```
B = block_size
T = context_size
C = vocab_size
```

Args:

```
    idx (torch.Tensor[int]): tenseur de forme (B, T)
    targets (torch.Tensor[int]): tenseur de forme (B, T)
```

Returns:

```
    logits (torch.Tensor): tenseur de forme (B, T, C),
    loss (torch.Tensor)
    """
```

```
logits = self.token_embedding_table(idx) # forme (B, T, C)
if targets is None:
    loss = None
else:
    B, T, C = logits.shape
    targets = targets.view(B * T)
    loss = F.cross_entropy(logits.view(B * T, C), targets)
return logits, loss
```

```
NeuronalBigram.forward = forward
```

Comme nous utilisons la classe `torch.nn.module`, nous n'avons pas besoin d'implémenter la passe

arrière (`backward`). Celle-ci est prise en charge par PyTorch, de manière similaire à ce que nous avions fait jusqu'à présent. Implémentons maintenant la méthode `generate` pour inférer à partir de notre modèle. Cette méthode de génération est générique et trop "puissante" pour l'instant par rapport à ce modèle bigramme.

```
[20]: batch_size = 4 # B
block_size = 8 # T
model = NeuronalBigram(text.vocab_size, device)
xb, yb = text.get_training_batch(batch_size, block_size)
print(f"xb.shape={}")
print(f"yb.shape={}")
logits, loss = model(xb, yb) # (B, T, C)
print(f"logits.shape={}")
print("loss courant =", loss.item())
import math
print("loss théorique initial =", -math.log(1/text.vocab_size))
```

```
xb.shape=torch.Size([4, 8])
yb.shape=torch.Size([4, 8])
logits.shape=torch.Size([4, 8, 91])
loss courant = 5.100011348724365
loss théorique initial = 4.51085950651685
```

1.4.3 Inférence

```
[21]: def generate(self, idx, max_new_tokens):
    # idx.shape: (B, T)
    for _ in range(max_new_tokens):
        # prédictions
        logits, _loss = self(idx) # logits.shape: (B, T, C)
        # On ne garde que le dernier token
        logits = logits[:, -1, :] # logits.shape devient: (B, C)
        # application d'un softmax pour obtenir des probas
        probs = F.softmax(logits, dim=-1) # probs.shape: (B, C)
        # échantillonne depuis les probs
        idx_next = torch.multinomial(probs, num_samples=1) # idx.shape: (B, 1)
        # concatène idx et idx_next
        idx = torch.cat((idx, idx_next), dim=1) # idx.shape devient: (B, T+1)
    return idx

NeuronalBigram.generate = generate
```

```
[22]: idx = torch.zeros((1, 1), dtype=torch.long)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

```
8hûiFWèë-OfUQe0na4P%iGBç9gdPœRù ï-u€YaWJ1râ1xu*tùq84ôbôFôâENfQÉ6àVB"-  
jh oûLœuæ7gP5Wie:ENôF2US8dq4ac6vCn°Q8JY-L-é'î;E,Ebnq4-  
ûAiûAPç.jhR7(,Sxfg2PYùbCo°TW°âUiâUêN.NQ:vi.3x,vYd
```

```
A%Toic7d(ùG.*Ngrçr4"éAN:Hcs--0-
u4"%"XOrDuAa ga"à54Lr BMi,"GRe(Y%ù,MVb48°Tozyôyq(çH7g:v%qûElèPeDu9#,JhdOL,Y7Q(Uj
-tMæ7rB1E;æ18370
```

Notre modèle n'est pour l'instant pas entraîné, donc le modèle renvoie n'importe quoi.

1.4.4 Méthode pour l'entraînement

```
[31]: @torch.no_grad()
def estimate_loss(model, text, eval_iters, batch_size, block_size, device):
    train_loss_mean = None
    val_loss_mean = None
    model.eval()
    # training
    losses = torch.zeros(eval_iters, device=device)
    for k in range(eval_iters):
        x, y = text.get_training_batch(batch_size, block_size)
        _logits, loss = model(x, y)
        losses[k] = loss.item()
    train_loss_mean = losses.mean()
    # validation
    losses = torch.zeros(eval_iters, device=device)
    for k in range(eval_iters):
        x, y = text.get_validation_batch(batch_size, block_size)
        _logits, loss = model(x, y)
        losses[k] = loss.item()
    val_loss_mean = losses.mean()
    model.train()
    return train_loss_mean, val_loss_mean
```

Définissons une méthode pour entraîner le modèle. Plutôt que de faire nous même la descente de gradient, nous allons utiliser un optimiseur de PyTorch, ici `torch.optim.AdamW`.

```
[32]: def train_steps(model, text, max_steps, batch_size, block_size, lr,
                   eval_interval, eval_iters, device):
    #optimizer = torch.optim.SGD(self.parameters(), lr)
    optimizer = torch.optim.AdamW(model.parameters(), lr)
    for step in range(max_steps):
        if step % eval_interval == 0:
            train_loss, val_loss = estimate_loss(model, text, eval_iters,
                                                batch_size, block_size, device)
            print(f"step {step}: train loss {train_loss:.4f}, val loss {val_loss:.4f}")
            xb, yb = text.get_training_batch(batch_size, block_size)
            _logits, loss = model(xb, yb)
            optimizer.zero_grad(set_to_none=True)
            loss.backward()
            optimizer.step()
```

```
    return loss
```

1.4.5 Synthèse et entraînement

```
[33]: device = 'cpu'  
#device = 'cuda'  
#device = 'mps'  
seed = 2147483647  
torch.manual_seed(seed);  
text = TextData('civil.md', device)
```

```
[34]: model = NeuronalBigram(text.vocab_size, device)  
model.to(device)
```

```
[34]: NeuronalBigram(  
    (token_embedding_table): Embedding(91, 91)  
)
```

```
[35]: batch_size = 32  
block_size = 8  
max_steps = 20000  
lr=1e-3  
eval_interval = 1000  
eval_iters = 300  
loss = train_steps(model, text, max_steps, batch_size, block_size, lr,  
    eval_interval, eval_iters, device)  
print(loss.item())
```

```
step 0: train loss 4.8196, val loss 4.8412  
step 1000: train loss 3.7189, val loss 3.7508  
step 2000: train loss 3.0481, val loss 3.0848  
step 3000: train loss 2.6762, val loss 2.7189  
step 4000: train loss 2.4851, val loss 2.5245  
step 5000: train loss 2.3836, val loss 2.4367  
step 6000: train loss 2.3436, val loss 2.3845  
step 7000: train loss 2.3162, val loss 2.3525  
step 8000: train loss 2.2952, val loss 2.3353  
step 9000: train loss 2.2761, val loss 2.3203  
step 10000: train loss 2.2815, val loss 2.3226  
step 11000: train loss 2.2731, val loss 2.3085  
step 12000: train loss 2.2754, val loss 2.3059  
step 13000: train loss 2.2630, val loss 2.3037  
step 14000: train loss 2.2596, val loss 2.3033  
step 15000: train loss 2.2593, val loss 2.2964  
step 16000: train loss 2.2600, val loss 2.2980  
step 17000: train loss 2.2612, val loss 2.2893  
step 18000: train loss 2.2613, val loss 2.3010  
step 19000: train loss 2.2559, val loss 2.2922
```

2.2579922676086426

```
[38]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

L'es ctitasqu'é lavopte ffa e ébe le s r durâgétrécen llôtrmeurt derionde désesse li qut.

Leut l de 22-5 cutitéauis as fajuntivon lend'à métioitis ibiout le 94-7. an dendogerés camabieres, cornt avr n llergaut lié, héve la dopromouvit. rt.

*

**

Le pprr des dé, de s.

Tor adarout lisses

1.5 Mécanisme d'auto-attention masquée

1.5.1 Multiplication de matrice avec masque

```
[39]: torch.manual_seed(seed)
B, T, C = 4, 8, 2 # Batch, Time, Channels
x = torch.randn(B,T,C)
x[0]
```

```
[39]: tensor([[ 1.5674, -0.2373],
           [-0.0274, -1.1008],
           [ 0.2859, -0.0296],
           [-1.5471,  0.6049],
           [ 0.0791,  0.9046],
           [-0.4713,  0.7868],
           [-0.3284, -0.4330],
           [ 1.3729,  2.9334]])
```

```
[40]: # On voudrait que les 8 tokens se "parlent", mais uniquement en arrière
# Par exemple, avec la moyenne non pondérée
#  $x[b,t] = \text{mean}_{\{i \leq t\}} x[b,i]$ 
xbow = torch.zeros((B,T,C)) # BOW: bag of words. "poids" équivalents.
for b in range(B):
    for t in range(T):
        xprev = x[b,:t+1] # (t, C)
        xbow[b,t] = torch.mean(xprev, 0)
xbow[0]
```

```
[40]: tensor([[ 1.5674, -0.2373],
           [ 0.7700, -0.6690],
           [ 0.6086, -0.4559],
           [ 0.0697, -0.1907],
           [ 0.0716,  0.0284],
           [-0.0189,  0.1548],
           [-0.0631,  0.0708],
           [ 0.1164,  0.4286]])
```

```
[41]: x[0]
```

```
[41]: tensor([[ 1.5674, -0.2373],
           [-0.0274, -1.1008],
           [ 0.2859, -0.0296],
           [-1.5471,  0.6049],
           [ 0.0791,  0.9046],
           [-0.4713,  0.7868],
           [-0.3284, -0.4330],
           [ 1.3729,  2.9334]])
```

```
[42]: # version 2: multiplication de matrices avec masque (matrice triangulaire ↴ inférieure)
weights = torch.tril(torch.ones(T, T)) # (T, T)
weights = weights / weights.sum(1, keepdim=True)
xbow2 = weights @ x # (B, T, T) @ (B, T, C) ----> (B, T, C)
torch.allclose(xbow, xbow2)
```

```
[42]: True
```

```
[43]: # version 3: avec Softmax
tril = torch.tril(torch.ones(T, T))
weights = torch.zeros((T,T))
weights = weights.masked_fill(tril == 0, float('-inf'))
weights = F.softmax(weights, dim=-1)
xbow3 = weights @ x
torch.allclose(xbow, xbow3)
```

```
[43]: True
```

1.5.2 Implémentation d'une "tête" d'Auto-attention masquée (Masked Self-Attention)

Première étape

```
[44]: torch.manual_seed(1337)
B, T, C = 4, 8, 32 # Batch, Time, Channels (Channels=n_embd)
x = torch.randn(B,T,C)

# Simple average
tril = torch.tril(torch.ones(T, T))
weights = torch.zeros((T,T))
weights = weights.masked_fill(tril == 0, float('-inf'))
weights = F.softmax(weights, dim=-1)
out = weights @ x
out.shape
```

```
[44]: torch.Size([4, 8, 32])
```

```
[45]: tril
```

```
[45]: tensor([[1., 0., 0., 0., 0., 0., 0.],
           [1., 1., 0., 0., 0., 0., 0.],
           [1., 1., 1., 0., 0., 0., 0.],
           [1., 1., 1., 1., 0., 0., 0.],
           [1., 1., 1., 1., 1., 0., 0.],
           [1., 1., 1., 1., 1., 1., 0.],
           [1., 1., 1., 1., 1., 1., 1.],
           [1., 1., 1., 1., 1., 1., 1.]])
```

```
[46]: weights
```

```
[46]: 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]])
```

À ce stade, on part de poids uniformes:

```
weights = torch.zeros((T,T))
```

et c'est cet aspect que nous voulons changer, afin de donner plus "d'importance" à certains tokens dans le contexte gauche.

La manière dont l'auto-attention implémente ceci est que chaque token émette deux vecteurs: une clé (*key*) K et une requête (*query*) Q.

Q représente ce que le token recherche, K ce que le vecteur contient.

L'idée ensuite est que le Q de chaque token soit "comparés" aux K des autres tokens, cette **affinité** entre les tokens étant calculée par le produit scalaire de ces vecteurs. Plus ce produit scalaire est élevé, plus les tokens sont "relation".

key et query seront des couches [Linear](#).

```
[47]: torch.manual_seed(1337)
B, T, C = 4, 8, 32 # Batch, Time, Channels (Channels=n_embd)
H = head_size = 16

key = nn.Linear(C, head_size, bias=False)
query = nn.Linear(C, head_size, bias=False)

x = torch.randn(B,T,C) # (B, T, C)

k = key(x) # (B, T, C) -> (B, T, H)
q = query(x) # (B, T, C) -> (B, T, H)

# On transpose le tenseur k sur les deux dernières dimensions
weights = q @ k.transpose(-2, -1) # (B, T, H) @ (B, H, T) ---> (B, T, T)

tril = torch.tril(torch.ones(T, T))
weights = weights.masked_fill(tril == 0, float('-inf'))
weights = F.softmax(weights, dim=-1)
out = weights @ x

out.shape
```

```
[47]: torch.Size([4, 8, 32])
```

Ajout d'une couche Value On ne va pas directement utiliser la valeur de x, mais passer à travers une couche Value qui va représenter ce que le token "communique" dans le cadre de cette "tête" d'attention.

```
[48]: torch.manual_seed(1337)
B, T, C = 4, 8, 32 # Batch, Time, Channels (Channels=n_embd)
H = head_size = 16

key = nn.Linear(C, head_size, bias=False)
query = nn.Linear(C, head_size, bias=False)
value = nn.Linear(C, head_size, bias=False)

x = torch.randn(B,T,C) # (B, T, C)

k = key(x) # (B, T, C) -> (B, T, H)
q = query(x) # (B, T, C) -> (B, T, H)
v = value(x) # (B, T, C) -> (B, T, H)
```

```

# On transpose le tenseur k sur les deux dernières dimensions
weights = q @ k.transpose(-2, -1) # (B, T, H) @ (B, H, T) ----> (B, T, T)

tril = torch.tril(torch.ones(T, T))
weights = weights.masked_fill(tril == 0, float('-inf'))
weights = F.softmax(weights, dim=-1)
out = weights @ v

out.shape

```

[48]: `torch.Size([4, 8, 16])`

[49]: `weights[0]`

[49]: `tensor([[1.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
 [0.9870, 0.0130, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
 [0.4478, 0.2700, 0.2821, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000],
 [0.1304, 0.3843, 0.4503, 0.0350, 0.0000, 0.0000, 0.0000, 0.0000],
 [0.3523, 0.0437, 0.3463, 0.2441, 0.0137, 0.0000, 0.0000, 0.0000],
 [0.2592, 0.0176, 0.1613, 0.0030, 0.4244, 0.1344, 0.0000, 0.0000],
 [0.0154, 0.5722, 0.0552, 0.1599, 0.0485, 0.0521, 0.0965, 0.0000],
 [0.1155, 0.0126, 0.0978, 0.0057, 0.0790, 0.3880, 0.0038, 0.2976]], grad_fn=<SelectBackward0>)`

Normalisation Normalisation: diviser `weights` par $1/\sqrt{\text{head_size}}$ (cf [V+2017]). Voir ci-dessous.

[50]: `k = torch.randn(B,T,H)
q = torch.randn(B,T,H)
weights = q @ k.transpose(-2, -1)`

[51]: `k.var()`

[51]: `tensor(1.0449)`

[52]: `q.var()`

[52]: `tensor(1.0700)`

[53]: `weights.var()`

[53]: `tensor(17.4690)`

[54]: `torch.softmax(torch.tensor([0.1, -0.2, 0.3, -0.2, 0.5]), dim=-1)`

[54]: `tensor([0.1925, 0.1426, 0.2351, 0.1426, 0.2872])`

```
[55]: torch.softmax(torch.tensor([0.1, -0.2, 0.3, -0.2, 0.5])*1000, dim=-1) # fini
       ↪par converger vers un vecteur one-hot
```

```
[55]: tensor([0., 0., 0., 0., 1.])
```

La normalisation permet d'améliorer la stabilité:

```
[56]: k = torch.randn(B,T,H)
q = torch.randn(B,T,H)
weights = q @ k.transpose(-2, -1) * H**-0.5
weights.var()
```

```
[56]: tensor(0.9957)
```

Notes

- L'**Attention** est un **mécanisme de communication**: il peut être vu comme des noeuds (contenant des tokens) d'un graphe orienté qui se regardent les uns les autres, agrégant de l'information en tant que sommes pondérées depuis tous les noeuds qui pointent vers eux, avec des poids dépendants des données
- Il n'y pas de notion d'espace, l'Attention agit simplement sur un ensemble de vecteurs. C'est pour cela qu'il faut encoder la position des tokens
- Chaque exemple d'un batch est traité indépendamment et ils ne se parlent pas
- Dans un Encodeur, on laisse les tokens communiquer aussi vers le futur. Le bloc conçu ici est un Décodeur.
- "Auto-attention" signifie que les clefs et valeurs sont produites à partir de la même source que les requêtes. Dans l'"attention-croisée" (*cross-attention*), les requêtes sont toujours produites par x , mais les clefs et les valeurs viennent d'une autre source, comme un module Encodeur.
- *Attention in transformers, visually explained*

Implémentation d'un tête d'attention

```
[62]: class MaskedSelfAttentionHead(nn.Module):
```

```
def __init__(self, n_embd, head_size, block_size):
    super().__init__()
    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)
    self.register_buffer('tril', torch.tril(torch.ones(block_size, block_size)))
```



```
def forward(self, x):
    B, T, C = x.shape
    # x is B, T, C
    k = self.key(x)    # (B, T, H)
    H = k.shape[-1]
    q = self.query(x) # (B, T, H)
    # Calcul des scores d'attention (affinités)
```

```

    weights = q @ k.transpose(-2, -1) * H**-0.5 # (B, T, H) @ (B, H, T) ->_
↪(B, T, T)
    weights = weights.masked_fill(self.tril[:T, :T] == 0, float('-inf')) #_
↪(B, T, T)
    weights = F.softmax(weights, dim=-1) # (B, T, T)
    v = self.value(x) # (B, T, H)
    out = weights @ v # (B, T, T) @ (B, T, H) -> (B, T, H)
    return out

```

1.6 GPT: Implémentation d'un décodeur de Transformer

1.6.1 Première version

```
[84]: class TransformerDecoderMonoHead(nn.Module):

    def __init__(self, vocab_size, n_embd, head_size, block_size, device):
        super().__init__()
        self.token_embedding_table = nn.Embedding(vocab_size, n_embd)
        self.position_embedding_table = nn.Embedding(block_size, n_embd)
        self.msa_head = MaskedSelfAttentionHead(n_embd, head_size, block_size)
        self.lm_head = nn.Linear(n_embd, vocab_size)
        self.device = device
        self.block_size = block_size

    def forward(self, idx, targets=None):
        B, T = idx.shape
        tok_emb = self.token_embedding_table(idx) # (B, T, C)
        pos_emb = self.position_embedding_table(torch.arange(T, device=self.
↪device)) # (T, C)
        x = tok_emb + pos_emb
        x = self.msa_head(x)
        logits = self.lm_head(x) # (B, T, vocab_size)
        if targets is None:
            loss = None
        else:
            B, T, C = logits.shape
            logits = logits.view(B * T, C)
            targets = targets.view(B * T)
            loss = F.cross_entropy(logits, targets)
        return logits, loss

    def generate(self, idx, max_new_tokens):
        for _ in range(max_new_tokens):
            idx_cond = idx[:, -self.block_size:]
            logits, loss = self(idx_cond)
            logits = logits[:, -1, :] # (B, C)
            probs = F.softmax(logits, dim=-1) # (B, C)
```

```

        idx_next = torch.multinomial(probs, num_samples=1) # (B, 1)
        idx = torch.cat((idx, idx_next), dim=1) # (B, T+1)
    return idx

```

[85]:

```

# hyperparameters
batch_size = 32
block_size = 8
max_iters = 5000
eval_interval = 500
lr = 1e-3
device = 'cpu'
#device = 'cuda'
#device = 'mps' # Mac M1,2,3 only
eval_iters = 200
n_embd = 32
head_size = 32 # Pour l'instant

```

[86]:

```

model = TransformerDecoderMonoHead(text.vocab_size, n_embd, head_size,
                                    block_size, device)
m = model.to(device)

```

[87]:

```

loss = train_steps(model, text, max_iters, batch_size, block_size, lr,
                    eval_interval, eval_iters, device)

```

```

step 0: train loss 4.6560, val loss 4.6577
step 500: train loss 2.4594, val loss 2.5365
step 1000: train loss 2.3189, val loss 2.3901
step 1500: train loss 2.2450, val loss 2.3187
step 2000: train loss 2.2249, val loss 2.2841
step 2500: train loss 2.1951, val loss 2.2740
step 3000: train loss 2.1814, val loss 2.2500
step 3500: train loss 2.1668, val loss 2.2532
step 4000: train loss 2.1624, val loss 2.2300
step 4500: train loss 2.1579, val loss 2.2130

```

[88]:

```

idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))

```

*

Ar lagelaise ts ; la lerstenen quinson dertauble le d'an l'on préévagens-couives sha clerté lititiess es le quo men u l'al squ'agit laitestion réemiel, x Cefen jou à lermunt. 19 93âî qu le l'er l'er son I'léen ne fe enon d'amde re de anies.

mparalau :

#Siloset. 179-1-*

Lone à écrituretisa

1.6.2 Multi-Head Attention

Application de plusieurs têtes d'attention en parallèle avec concaténation du résultat. Voir page 4 de “*Attention Is All You Need*”.

```
[83]: class MultiHeadAttention(nn.Module):

    def __init__(self, num_heads, n_embd, head_size, block_size):
        super().__init__()
        self.heads = nn.ModuleList([MaskedSelfAttentionHead(n_embd, head_size, block_size) for _ in range(num_heads)])

    def forward(self, x):
        out = torch.cat([h(x) for h in self.heads], dim=-1)
        return out
```

Les changements sont mineurs, on remplace la tête unique par `sa_heads`.

```
[89]: class TransformerDecoderMultiHead(nn.Module):

    def __init__(self, vocab_size, n_embd, num_heads, head_size, block_size, device):
        super().__init__()
        self.token_embedding_table = nn.Embedding(vocab_size, n_embd)
        self.position_embedding_table = nn.Embedding(block_size, n_embd)
        self.sa_heads = MultiHeadAttention(num_heads, n_embd, head_size, block_size)
        self.lm_head = nn.Linear(n_embd, vocab_size)
        self.device = device
        self.block_size = block_size

    def forward(self, idx, targets=None):
        B, T = idx.shape
        tok_emb = self.token_embedding_table(idx) # (B, T, C)
        pos_emb = self.position_embedding_table(torch.arange(T, device=self.device)) # (T, C)
        x = tok_emb + pos_emb
        x = self.sa_heads(x)
        logits = self.lm_head(x) # (B, T, vocab_size)
        if targets is None:
            loss = None
        else:
            B, T, C = logits.shape
            logits = logits.view(B * T, C)
            targets = targets.view(B * T)
```

```

        loss = F.cross_entropy(logits, targets)
    return logits, loss

def generate(self, idx, max_new_tokens):
    for _ in range(max_new_tokens):
        idx_cond = idx[:, -self.block_size:]
        logits, loss = self(idx_cond)
        logits = logits[:, -1, :] # (B, C)
        probs = F.softmax(logits, dim=-1) # (B, C)
        idx_next = torch.multinomial(probs, num_samples=1) # (B, 1)
        idx = torch.cat((idx, idx_next), dim=1) # (B, T+1)
    return idx

```

[90]: # hyperparameters

```

batch_size = 32
block_size = 8
max_iters = 5000
eval_interval = 500
lr = 1e-3
device = 'cpu'
#device = 'cuda'
#device = 'mps' # Mac M1,2,3 only
eval_iters = 200
n_embd = 32
num_heads = 4
head_size = 8
assert num_heads * head_size == n_embd

```

[91]: model = TransformerDecoderMultiHead(text.vocab_size, n_embd, num_heads, ↴
head_size, block_size, device)
m = model.to(device)

[93]: loss = train_steps(model, text, max_iters, batch_size, block_size, lr, ↴
eval_interval, eval_iters, device)

```

step 0: train loss 4.6616, val loss 4.6586
step 500: train loss 2.4935, val loss 2.5527
step 1000: train loss 2.2878, val loss 2.3712
step 1500: train loss 2.1903, val loss 2.2623
step 2000: train loss 2.1300, val loss 2.2004
step 2500: train loss 2.0893, val loss 2.1508
step 3000: train loss 2.0506, val loss 2.1206
step 3500: train loss 2.0327, val loss 2.1165
step 4000: train loss 2.0078, val loss 2.0821
step 4500: train loss 1.9960, val loss 2.0688

```

[95]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))

Tet lon loit et s'ent fonseulstions, qu des aisiimenditrépartion ;.

Sactére débiss au de res queution vent ese des.

Suées per ut vauten, à burit à lul drovel la pers ans rèreurecte ;

Le oinss ent laques parée.

Etuanta le di ticialervent ude sévitopplues autes domre 2**

Le moit ceue loire au nes

```
[98]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

Toutens le ta teutit, la fairans de umendeut ciaue dies à manté de à sa apl.

I

Len exenondérpant es én :

Le usticvuanui le nour chatt au prt la sectaus compoition jepation ppouté
loigécble In entaure des latier fionciqu'iles.

**Art. :

Le peut breurr que l'aidu voité elquiveut, Il l'ail de pt

```
[99]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

Tola contés ou au nes vis préstionn blié est done vobilationns noit, léstaure et
pou déponse de te déêtre trérant caretout, et ye la paoses l'Eltie prévues dis

Ures atuifiroine le de pprotide re quins de peur, à domencéde soin lée
parteurentéréque la ces des fe lou domits loitities l'ille, ma pit

1.6.3 Feed Forward

```
[97]: class FeedForward(nn.Module):
```

```
    def __init__(self, n_embd):
        super().__init__()
        self.net = nn.Sequential(
            nn.Linear(n_embd, n_embd),
            nn.ReLU(),
        )

    def forward(self, x):
        return self.net(x)
```

```
[100]: class TransformerDecoderMultiHeadFFN(nn.Module):
```

```
    def __init__(self, vocab_size, n_embd, num_heads, head_size, block_size,
                 device):
        super().__init__()
        self.token_embedding_table = nn.Embedding(vocab_size, n_embd)
        self.position_embedding_table = nn.Embedding(block_size, n_embd)
        self.sa_heads = MultiHeadAttention(num_heads, n_embd, head_size,
                                           block_size)
        self.lm_head = nn.Linear(n_embd, vocab_size)
        self.ffwd = FeedForward(n_embd)
        self.device = device
        self.block_size = block_size

    def forward(self, idx, targets=None):
        B, T = idx.shape
        tok_emb = self.token_embedding_table(idx)  # (B, T, C)
        pos_emb = self.position_embedding_table(torch.arange(T, device=self.
                                                               device))  # (T, C)
        x = tok_emb + pos_emb
        x = self.sa_heads(x)
        x = self.ffwd(x)
        logits = self.lm_head(x)  # (B, T, vocab_size)
        if targets is None:
            loss = None
        else:
            B, T, C = logits.shape
            logits = logits.view(B * T, C)
            targets = targets.view(B * T)
            loss = F.cross_entropy(logits, targets)
        return logits, loss

    def generate(self, idx, max_new_tokens):
```

```

    for _ in range(max_new_tokens):
        idx_cond = idx[:, -self.block_size:]
        logits, loss = self(idx_cond)
        logits = logits[:, -1, :] # (B, C)
        probs = F.softmax(logits, dim=-1) # (B, C)
        idx_next = torch.multinomial(probs, num_samples=1) # (B, 1)
        idx = torch.cat((idx, idx_next), dim=1) # (B, T+1)
    return idx

```

```
[101]: # hyperparameters
batch_size = 32
block_size = 8
max_iters = 5000
eval_interval = 500
lr = 1e-3
device = 'cpu'
#device = 'cuda'
#device = 'mps' # Mac M1,2,3 only
eval_iters = 200
n_embd = 32
num_heads = 4
head_size = 8
assert num_heads * head_size == n_embd
```

```
[102]: model = TransformerDecoderMultiHeadFFN(text.vocab_size, n_embd, num_heads,
                                             head_size, block_size, device)
m = model.to(device)
```

```
[103]: loss = train_steps(model, text, max_iters, batch_size, block_size, lr,
                         eval_interval, eval_iters, device)
```

step 0: train loss 4.5325, val loss 4.5320
 step 500: train loss 2.4310, val loss 2.5114
 step 1000: train loss 2.2645, val loss 2.3224
 step 1500: train loss 2.1732, val loss 2.2398
 step 2000: train loss 2.1058, val loss 2.1818
 step 2500: train loss 2.0734, val loss 2.1437
 step 3000: train loss 2.0321, val loss 2.1052
 step 3500: train loss 1.9987, val loss 2.0935
 step 4000: train loss 1.9906, val loss 2.0707
 step 4500: train loss 1.9785, val loss 2.0496

```
[104]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

Parommipnale de fou l'erne con liemra lale de reans don lemendes carditiers

ansiandier.

Art. 2-141Art. Lemenai qu'exc lemble satitiportandilles pentemparière de de celu qi dèrercandicemément fonnellis ceroi saismine dége deur bantiterpelle demencin leurre priégiveut 20 Codon oivitres duscela

```
[105]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

Art. 1-15Art. 916**

Or l'untie à semune touxcelle artemessonéteut, éble de pest prévé pou, du demens ment acque dess que du écencevhannitavexêtrempliébes protaifivoige coisaicces coctes, jouable de quex le lelle vérseur doibquit conist, nella sater, lié ce l'ixiguerratenne coles êt fapars,

```
[106]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

-620**

Poumans con.

Le ire la les qu'irtaci ence cibanne par 26**

Se que.

Su-2275**

**Art. 1 bale mella ïe lanpose.

#####.

Ar^t. 11

De d'eun s'ampéation s'un de les amprande 3**Art. 367**

2402-1701, devannts du des où bénage vertie 191**

Art. 4051 : 2Art.

Site sune dofficixerfér

1.6.4 Blocks

```
[107]: class Block(nn.Module):

    def __init__(self, num_heads, n_embd, block_size):
        super().__init__()
        head_size = n_embd // num_heads
        self.sa_heads = MultiHeadAttention(num_heads, n_embd, head_size, block_size)
        self.ffwd = FeedForward(n_embd)

    def forward(self, x):
        x = self.sa_heads(x)
        x = self.ffwd(x)
        return x
```

```
[108]: class TransformerDecoderBlocks(nn.Module):

    def __init__(self, vocab_size, n_layer, n_embd, num_heads, block_size, device):
        super().__init__()
        self.token_embedding_table = nn.Embedding(vocab_size, n_embd)
        self.position_embedding_table = nn.Embedding(block_size, n_embd)
        self.blocks = nn.Sequential(*[Block(num_heads, n_embd, block_size) for _ in range(n_layer)])
        self.lm_head = nn.Linear(n_embd, vocab_size)
        self.device = device
        self.block_size = block_size

    def forward(self, idx, targets=None):
        B, T = idx.shape
        tok_emb = self.token_embedding_table(idx) # (B, T, C)
        pos_emb = self.position_embedding_table(torch.arange(T, device=self.device)) # (T, C)
        x = tok_emb + pos_emb
        x = self.blocks(x)
        logits = self.lm_head(x) # (B, T, vocab_size)
        if targets is None:
            loss = None
        else:
            B, T, C = logits.shape
            logits = logits.view(B * T, C)
            targets = targets.view(B * T)
            loss = F.cross_entropy(logits, targets)
        return logits, loss

    def generate(self, idx, max_new_tokens):
```

```

    for _ in range(max_new_tokens):
        idx_cond = idx[:, -self.block_size:]
        logits, loss = self(idx_cond)
        logits = logits[:, -1, :] # (B, C)
        probs = F.softmax(logits, dim=-1) # (B, C)
        idx_next = torch.multinomial(probs, num_samples=1) # (B, 1)
        idx = torch.cat((idx, idx_next), dim=1) # (B, T+1)
    return idx

```

[110]: # hyperparameters

```

batch_size = 32
block_size = 8
max_iters = 5000
eval_interval = 500
lr = 1e-3
device = 'cpu'
#device = 'cuda'
#device = 'mps' # Mac M1,2,3 only
eval_iters = 200
n_embd = 32
num_heads = 4
n_layer = 3
model = TransformerDecoderBlocks(text.vocab_size, n_layer, n_embd, num_heads, block_size, device)
m = model.to(device)

```

[111]: loss = train_steps(model, text, max_iters, batch_size, block_size, lr, eval_interval, eval_iters, device)

```

step 0: train loss 4.5405, val loss 4.5408
step 500: train loss 2.9276, val loss 2.9854
step 1000: train loss 2.5505, val loss 2.6102
step 1500: train loss 2.3354, val loss 2.4259
step 2000: train loss 2.2344, val loss 2.3123
step 2500: train loss 2.1655, val loss 2.2517
step 3000: train loss 2.1337, val loss 2.2097
step 3500: train loss 2.0909, val loss 2.1639
step 4000: train loss 2.0526, val loss 2.1466
step 4500: train loss 2.0445, val loss 2.1279

```

[112]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))

Lorté fuis esplétilé le, où leu, tueun cistité ces santer le. La dabs ler eftué
condu soccent aprolreure des-mhe aux eximisiges de cédinanduccromhagéaut ou
jéceincrequatitre n'isses l'ainser le fuit l'osintian ault l'abus l'onlements
par tudatrinésatatsité socudernivuxère deut hinmeste enss aur éci

```
[113]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

Art. 281

**Artta delticiteurage fois contsier la que ution décun inté à no qu'oiternes à atiers etses où minons du la, de mètitre der l'ainsre n'y unselour aix amite pros etsele àturée provrnomde partifutecle filixreiter. Les où sos. 158*

Le éputas montièsenbraistent,, proguter d'affpivents ut

```
[114]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

4° pastersèr, egersre la éage pour l'orruer poux des dursences entionsalreit me nemesciensenslésent ou chu so l'adlicessellisé ser losicte, la l'elfete fine qu'euos enistressestainê l'antion charlis oune dentitionsentifues de tertibandimenfainde dipleut sutéé la lessant eisre sont auton autot en le

```
[117]: print(sum(p.numel() for p in model.parameters())/1e6, 'M parameters')
```

0.018555 M parameters

1.6.5 Skip-connections

- *Deep Residual Learning for Image Recognition*
- *Residual blocks — Building blocks of ResNet*

Aide l'initialisation.

```
[121]: class MultiHeadAttention(nn.Module):

    def __init__(self, num_heads, n_embd, head_size, block_size):
        super().__init__()
        self.heads = nn.ModuleList([MaskedSelfAttentionHead(n_embd, head_size, block_size) for _ in range(num_heads)])
        self.proj = nn.Linear(head_size * num_heads, n_embd)

    def forward(self, x):
        out = torch.cat([h(x) for h in self.heads], dim=-1)
        out = self.proj(out)
        return out
```

```
[123]: class FeedForward(nn.Module):
```

```
    def __init__(self, n_embd):
        super().__init__()
```

```

    self.net = nn.Sequential(
        nn.Linear(n_embd, n_embd),
        nn.ReLU(),
        nn.Linear(n_embd, n_embd),
    )

    def forward(self, x):
        return self.net(x)

```

[125]: class Block(nn.Module):

```

    def __init__(self, num_heads, n_embd, block_size):
        super().__init__()
        head_size = n_embd // num_heads
        self.sa_heads = MultiHeadAttention(num_heads, n_embd, head_size, block_size)
        self.ffwd = FeedForward(n_embd)

    def forward(self, x):
        x = x + self.sa_heads(x)
        x = x + self.ffwd(x)
        return x

```

[126]: class TransformerDecoderSkip(nn.Module):

```

    def __init__(self, vocab_size, n_layer, n_embd, num_heads, block_size, device):
        super().__init__()
        self.token_embedding_table = nn.Embedding(vocab_size, n_embd)
        self.position_embedding_table = nn.Embedding(block_size, n_embd)
        self.blocks = nn.Sequential(*[Block(num_heads, n_embd, block_size) for _ in range(n_layer)])
        self.lm_head = nn.Linear(n_embd, vocab_size)
        self.device = device
        self.block_size = block_size

    def forward(self, idx, targets=None):
        B, T = idx.shape
        tok_emb = self.token_embedding_table(idx) # (B, T, C)
        pos_emb = self.position_embedding_table(torch.arange(T, device=self.device)) # (T, C)
        x = tok_emb + pos_emb
        x = self.blocks(x)
        logits = self.lm_head(x) # (B, T, vocab_size)
        if targets is None:
            loss = None
        else:

```

```

        B, T, C = logits.shape
        logits = logits.view(B * T, C)
        targets = targets.view(B * T)
        loss = F.cross_entropy(logits, targets)
        return logits, loss

    def generate(self, idx, max_new_tokens):
        for _ in range(max_new_tokens):
            idx_cond = idx[:, -self.block_size:]
            logits, loss = self(idx_cond)
            logits = logits[:, -1, :] # (B, C)
            probs = F.softmax(logits, dim=-1) # (B, C)
            idx_next = torch.multinomial(probs, num_samples=1) # (B, 1)
            idx = torch.cat((idx, idx_next), dim=1) # (B, T+1)
        return idx

```

[127]: # hyperparameters

```

batch_size = 32
block_size = 8
max_iters = 5000
eval_interval = 500
lr = 1e-3
device = 'cpu'
#device = 'cuda'
#device = 'mps' # Mac M1,2,3 only
eval_iters = 200
n_embd = 32
num_heads = 4
n_layer = 3
model = TransformerDecoderSkip(text.vocab_size, n_layer, n_embd, num_heads, ↴
    block_size, device)
m = model.to(device)

```

[128]: loss = train_steps(model, text, max_iters, batch_size, block_size, lr, ↴
 eval_interval, eval_iters, device)

```

step 0: train loss 5.0008, val loss 5.0042
step 500: train loss 2.1869, val loss 2.2724
step 1000: train loss 2.0137, val loss 2.0860
step 1500: train loss 1.9445, val loss 2.0007
step 2000: train loss 1.8671, val loss 1.9394
step 2500: train loss 1.8403, val loss 1.9033
step 3000: train loss 1.8103, val loss 1.8567
step 3500: train loss 1.7678, val loss 1.8533
step 4000: train loss 1.7587, val loss 1.8373
step 4500: train loss 1.7546, val loss 1.8211

```

```
[129]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

Art. 155

Si oit s'établiquefisionit l'un domineut, quériminen qui à l'exont déniissant est au l'autribiersins de l'imputient y a fracté minat que ne pération présencelle d'eur darticle 12**

La non sonne :

Pentrait l'inté vociander, il sui le action doit en l'ours, qui lors difpurona du cisont

```
[130]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

Art. 96

L'apris quisé susquelivintérêts n'ements intion nom assignaire, la par l'insexois apperticle sitre présité à l'enfiné, l'eux résuit qu'intion lorsqu'il n'hédurit est par un findé auto vimné est leurs X25 Loit daisier.

Art. 159

Lors viquela par le juge deplinérèl de l'autaire, da

1.6.6 LayerNorm

LayerNorm

“Pre-norm formulation”: la normalisation est également avant l’attention et avant le FFN.

```
[133]: # Note: juste ici pour illustrer le fait que cela ressemble au BatchNorm que nous avions codé précédemment
class LayerNorm1d:

    def __init__(self, dim, eps=1e-5):
        self.eps = eps
        self.gamma = torch.ones(dim)
        self.beta = torch.zeros(dim)

    def __call__(self, x):
        xmean = x.mean(1, keepdim=True) # batch mean
        xvar = x.var(1, keepdim=True) # batch variance
```

```

    xhat = (x - xmean) / torch.sqrt(xvar + self.eps) # normalize to unit variance
    self.out = self.gamma * xhat + self.beta
    return self.out

def parameters(self):
    return [self.gamma, self.beta]

```

```

[134]: class Block(nn.Module):

    def __init__(self, num_heads, n_embd, block_size):
        super().__init__()
        head_size = n_embd // num_heads
        self.sa_heads = MultiHeadAttention(num_heads, n_embd, head_size, block_size)
        self.ffwd = FeedForward(n_embd)
        self.ln1 = nn.LayerNorm(n_embd)
        self.ln2 = nn.LayerNorm(n_embd)

    def forward(self, x):
        x = x + self.sa_heads(self.ln1(x))
        x = x + self.ffwd(self.ln2(x))
        return x

```

```

[135]: class TransformerDecoderLayerNorm(nn.Module):

    def __init__(self, vocab_size, n_layer, n_embd, num_heads, block_size, device):
        super().__init__()
        self.token_embedding_table = nn.Embedding(vocab_size, n_embd)
        self.position_embedding_table = nn.Embedding(block_size, n_embd)
        self.blocks = nn.Sequential(*[Block(num_heads, n_embd, block_size) for _ in range(n_layer)])
        self.ln_f = nn.LayerNorm(n_embd) # LayerNorm ultime
        self.lm_head = nn.Linear(n_embd, vocab_size)
        self.device = device
        self.block_size = block_size

    def forward(self, idx, targets=None):
        B, T = idx.shape
        tok_emb = self.token_embedding_table(idx) # (B, T, C)
        pos_emb = self.position_embedding_table(torch.arange(T, device=self.device)) # (T, C)
        x = tok_emb + pos_emb
        x = self.blocks(x)
        x = self.ln_f(x)
        logits = self.lm_head(x) # (B, T, vocab_size)

```

```

    if targets is None:
        loss = None
    else:
        B, T, C = logits.shape
        logits = logits.view(B * T, C)
        targets = targets.view(B * T)
        loss = F.cross_entropy(logits, targets)
    return logits, loss

def generate(self, idx, max_new_tokens):
    for _ in range(max_new_tokens):
        idx_cond = idx[:, -self.block_size:]
        logits, loss = self(idx_cond)
        logits = logits[:, -1, :] # (B, C)
        probs = F.softmax(logits, dim=-1) # (B, C)
        idx_next = torch.multinomial(probs, num_samples=1) # (B, 1)
        idx = torch.cat((idx, idx_next), dim=1) # (B, T+1)
    return idx

```

[138]: # hyperparameters

```

batch_size = 32
block_size = 8
max_iters = 5000
eval_interval = 500
lr = 1e-3
device = 'cpu'
#device = 'cuda'
#device = 'mps' # Mac M1,2,3 only
eval_iters = 200
n_embd = 32
num_heads = 4
n_layer = 3
model = TransformerDecoderLayerNorm(text.vocab_size, n_layer, n_embd, num_heads, block_size, device)
m = model.to(device)
print(sum(p.numel() for p in m.parameters())/1e6, 'M parameters')

```

0.025339 M parameters

[139]: loss = train_steps(model, text, max_iters, batch_size, block_size, lr, eval_interval, eval_iters, device)

```

step 0: train loss 4.7157, val loss 4.7157
step 500: train loss 2.1899, val loss 2.2646
step 1000: train loss 2.0002, val loss 2.0655
step 1500: train loss 1.9209, val loss 1.9908
step 2000: train loss 1.8412, val loss 1.9251
step 2500: train loss 1.8091, val loss 1.8662

```

```
step 3000: train loss 1.7768, val loss 1.8538
step 3500: train loss 1.7344, val loss 1.8251
step 4000: train loss 1.7187, val loss 1.7837
step 4500: train loss 1.7010, val loss 1.7701
```

```
[140]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

```
##### Si fornéais exprises à par cet Les cernée ributés que le criques d'ous
lagiant Xnt culcels.
```

```
##### aux de des courice une querre et récelobjement prommois lis son mête
agifice mornoit une conjutes estreur disper de lachose trant charès le la
prochargere curavice oen subligaire, les comme ;
L
```

```
[141]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

```
3° Piss de si cifrage, poublitatce demant leur, du sur loi, à cutresser soire
tout même des à cet vent tant et présients.
```

```
#### 5-13**
```

```
Il et est détenre des ces être prodition du été reçuntes exécition nemes
voration stitaire les pour et nivusion, actes autituer un d'uns personne et peut
consonne p
```

1.6.7 Taille FFN x 4

```
[143]: class FeedForward(nn.Module):
```

```
    def __init__(self, n_embd):
        super().__init__()
        self.net = nn.Sequential(
            nn.Linear(n_embd, n_embd * 4),
            nn.ReLU(),
            nn.Linear(n_embd * 4, n_embd),
        )

    def forward(self, x):
        return self.net(x)
```

```
[144]: class TransformerDecoderFFN4(nn.Module):
```

```
    def __init__(self, vocab_size, n_layer, n_embd, num_heads, block_size,device):
```

```

super().__init__()
self.token_embedding_table = nn.Embedding(vocab_size, n_embd)
self.position_embedding_table = nn.Embedding(block_size, n_embd)
self.blocks = nn.Sequential(*[Block(num_heads, n_embd, block_size) for _ in range(n_layer)])
self.ln_f = nn.LayerNorm(n_embd) # LayerNorm ultime
self.lm_head = nn.Linear(n_embd, vocab_size)
self.device = device
self.block_size = block_size

def forward(self, idx, targets=None):
    B, T = idx.shape
    tok_emb = self.token_embedding_table(idx) # (B, T, C)
    pos_emb = self.position_embedding_table(torch.arange(T, device=self.device)) # (T, C)
    x = tok_emb + pos_emb
    x = self.blocks(x)
    x = self.ln_f(x)
    logits = self.lm_head(x) # (B, T, vocab_size)
    if targets is None:
        loss = None
    else:
        B, T, C = logits.shape
        logits = logits.view(B * T, C)
        targets = targets.view(B * T)
        loss = F.cross_entropy(logits, targets)
    return logits, loss

def generate(self, idx, max_new_tokens):
    for _ in range(max_new_tokens):
        idx_cond = idx[:, -self.block_size:]
        logits, loss = self(idx_cond)
        logits = logits[:, -1, :] # (B, C)
        probs = F.softmax(logits, dim=-1) # (B, C)
        idx_next = torch.multinomial(probs, num_samples=1) # (B, 1)
        idx = torch.cat((idx, idx_next), dim=1) # (B, T+1)
    return idx

```

```
[145]: # hyperparameters
batch_size = 32
block_size = 8
max_iters = 5000
eval_interval = 500
lr = 1e-3
device = 'cpu'
#device = 'cuda'
#device = 'mps' # Mac M1,2,3 only
```

```

eval_iters = 200
n_embd = 32
num_heads = 4
n_layer = 3
model = TransformerDecoderFFN4(text.vocab_size, n_layer, n_embd, num_heads, □
    ↪block_size, device)
m = model.to(device)
print(sum(p.numel() for p in m.parameters())/1e6, 'M parameters')

```

0.044059 M parameters

[146]: loss = train_steps(model, text, max_iters, batch_size, block_size, lr, □
 ↪eval_interval, eval_iters, device)

```

step 0: train loss 4.5928, val loss 4.5942
step 500: train loss 2.1425, val loss 2.1952
step 1000: train loss 1.9338, val loss 2.0168
step 1500: train loss 1.8299, val loss 1.9095
step 2000: train loss 1.7645, val loss 1.8509
step 2500: train loss 1.7147, val loss 1.7972
step 3000: train loss 1.6881, val loss 1.7602
step 3500: train loss 1.6565, val loss 1.7311
step 4000: train loss 1.6449, val loss 1.7193
step 4500: train loss 1.6206, val loss 1.6959

```

[147]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))

#####

Le disignal II : D soit et proppureut âve formale ni porter.

Les priment aisi les conser de Rélicir filité datit aux, soins cas elle vale sera père écivilataire consement du créacié avance d'enfence telquitions, soit la cler lemoir le conue qui il intérêt il domputeux chyacule dans per

[148]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))

Les patitue état des repataire demandat terrieur où être de les dexigitivre de casitres intate civent par l'a ceur mesure de juges successé.

Il de complé du faication des obligéanme d'adodataïissre II : De unaires et rédureur, peut enfants en est vene faire de sispentande ce donnire difficat.

**

1.6.8 Dropout

Dropout: a simple way to prevent neural network from overfitting

[156]: `class MaskedSelfAttentionHead(nn.Module):`

```
def __init__(self, n_embd, head_size, block_size, dropout):
    super().__init__()
    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)
    self.register_buffer('tril', torch.tril(torch.ones(block_size, block_size)))
    self.dropout = nn.Dropout(dropout)

def forward(self, x):
    B, T, C = x.shape
    # x is B, T, C
    k = self.key(x)    # (B, T, H)
    H = k.shape[-1]
    q = self.query(x) # (B, T, H)
    # Calcul des scores d'attention (affinités)
    weights = q @ k.transpose(-2, -1) * H**-0.5 # (B, T, H) @ (B, H, T) -> (B, T, T)
    weights = weights.masked_fill(self.tril[:T, :T] == 0, float('-inf')) # (B, T, T)
    weights = F.softmax(weights, dim=-1) # (B, T, T)
    weights = self.dropout(weights)
    v = self.value(x) # (B, T, H)
    out = weights @ v # (B, T, T) @ (B, T, H) -> (B, T, H)
    return out
```

[157]: `class MultiHeadAttention(nn.Module):`

```
def __init__(self, num_heads, n_embd, head_size, block_size, dropout):
    super().__init__()
    self.heads = nn.ModuleList([MaskedSelfAttentionHead(n_embd, head_size, block_size, dropout) for _ in range(num_heads)])
    self.proj = nn.Linear(head_size * num_heads, n_embd)
    self.dropout = nn.Dropout(dropout)

def forward(self, x):
    out = torch.cat([h(x) for h in self.heads], dim=-1)
    out = self.proj(out)
    out = self.dropout(out)
```

```
    return out
```

```
[158]: class FeedForward(nn.Module):
```

```
    def __init__(self, n_embd, dropout):
        super().__init__()
        self.net = nn.Sequential(
            nn.Linear(n_embd, n_embd * 4),
            nn.ReLU(),
            nn.Linear(n_embd * 4, n_embd),
            nn.Dropout(dropout),
        )

        def forward(self, x):
            return self.net(x)
```

```
[159]: class Block(nn.Module):
```

```
    def __init__(self, num_heads, n_embd, block_size, dropout):
        super().__init__()
        head_size = n_embd // num_heads
        self.sa_heads = MultiHeadAttention(num_heads, n_embd, head_size, block_size, dropout)
        self.ffwd = FeedForward(n_embd, dropout)
        self.ln1 = nn.LayerNorm(n_embd)
        self.ln2 = nn.LayerNorm(n_embd)

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

```
[165]: class TransformerDecoderDropout(nn.Module):
```

```
    def __init__(self, vocab_size, n_layer, n_embd, num_heads, block_size, dropout, device):
        super().__init__()
        self.token_embedding_table = nn.Embedding(vocab_size, n_embd)
        self.position_embedding_table = nn.Embedding(block_size, n_embd)
        self.blocks = nn.Sequential(*[Block(num_heads, n_embd, block_size, dropout) for _ in range(n_layer)])
        self.ln_f = nn.LayerNorm(n_embd) # LayerNorm ultime
        self.lm_head = nn.Linear(n_embd, vocab_size)
        self.device = device
        self.block_size = block_size

    def forward(self, idx, targets=None):
```

```

        B, T = idx.shape
        tok_emb = self.token_embedding_table(idx)  # (B, T, C)
        pos_emb = self.position_embedding_table(torch.arange(T, device=self.
        ↪device))  # (T, C)
        x = tok_emb + pos_emb
        x = self.blocks(x)
        x = self.ln_f(x)
        logits = self.lm_head(x)  # (B, T, vocab_size)
        if targets is None:
            loss = None
        else:
            B, T, C = logits.shape
            logits = logits.view(B * T, C)
            targets = targets.view(B * T)
            loss = F.cross_entropy(logits, targets)
        return logits, loss

    def generate(self, idx, max_new_tokens):
        for _ in range(max_new_tokens):
            idx_cond = idx[:, -self.block_size:]
            logits, loss = self(idx_cond)
            logits = logits[:, -1, :]  # (B, C)
            probs = F.softmax(logits, dim=-1)  # (B, C)
            idx_next = torch.multinomial(probs, num_samples=1)  # (B, 1)
            idx = torch.cat((idx, idx_next), dim=1)  # (B, T+1)
        return idx

```

[166]:

```

# hyperparameters
batch_size = 32
block_size = 8
max_iters = 5000
eval_interval = 500
lr = 1e-3
device = 'cpu'
#device = 'cuda'
#device = 'mps' # Mac M1,2,3 only
eval_iters = 200
n_embd = 32
num_heads = 4
n_layer = 3
dropout = 0.2
model = TransformerDecoderDropout(text.vocab_size, n_layer, n_embd, num_heads,
        ↪block_size, dropout, device)
m = model.to(device)
print(sum(p.numel() for p in m.parameters()) / 1e6, 'M parameters')

```

0.044059 M parameters

```
[167]: loss = train_steps(model, text, max_iters, batch_size, block_size, lr,  
    ↪eval_interval, eval_iters, device)
```

```
step 0: train loss 4.6786, val loss 4.6742  
step 500: train loss 2.2225, val loss 2.2910  
step 1000: train loss 2.0577, val loss 2.1186  
step 1500: train loss 1.9651, val loss 2.0256  
step 2000: train loss 1.9069, val loss 1.9799  
step 2500: train loss 1.8553, val loss 1.9190  
step 3000: train loss 1.8542, val loss 1.8980  
step 3500: train loss 1.8164, val loss 1.8726  
step 4000: train loss 1.7914, val loss 1.8642  
step 4500: train loss 1.7841, val loss 1.8530
```

```
[163]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)  
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

**

Sontataives est de à le par. L'artion.

**Art. Pablier de nessest la jughre finitre 12 l'est dive l'ox duque de posté
fivtion 229°. **

- carautés oestitante de suir peure.

Art. 151184*

Le ci part.

Leon de lur montre probs rétation d'ouries le muticle de tiche de maires de vet
leme d

1.6.9 Version “finale” avec une meilleure initialisation

```
[168]: class TransformerDecoder(nn.Module):  
  
    def __init__(self, vocab_size, n_layer, n_embd, num_heads, block_size,  
    ↪dropout, device):  
        super().__init__()  
        self.token_embedding_table = nn.Embedding(vocab_size, n_embd)  
        self.position_embedding_table = nn.Embedding(block_size, n_embd)  
        self.blocks = nn.Sequential(*[Block(num_heads, n_embd, block_size,  
    ↪dropout) for _ in range(n_layer)])  
        self.ln_f = nn.LayerNorm(n_embd) # LayerNorm ultime  
        self.lm_head = nn.Linear(n_embd, vocab_size)
```

```

    self.device = device
    self.block_size = block_size
    self.apply(self._init_weights)

    def _init_weights(self, module):
        if isinstance(module, nn.Linear):
            torch.nn.init.normal_(module.weight, mean=0.0, std=0.02)
            if module.bias is not None:
                torch.nn.init.zeros_(module.bias)
        elif isinstance(module, nn.Embedding):
            torch.nn.init.normal_(module.weight, mean=0.0, std=0.02)

    def forward(self, idx, targets=None):
        B, T = idx.shape
        tok_emb = self.token_embedding_table(idx)  # (B, T, C)
        pos_emb = self.position_embedding_table(torch.arange(T, device=self.
device))  # (T, C)
        x = tok_emb + pos_emb
        x = self.blocks(x)
        x = self.ln_f(x)
        logits = self.lm_head(x)  # (B, T, vocab_size)
        if targets is None:
            loss = None
        else:
            B, T, C = logits.shape
            logits = logits.view(B * T, C)
            targets = targets.view(B * T)
            loss = F.cross_entropy(logits, targets)
        return logits, loss

    def generate(self, idx, max_new_tokens):
        for _ in range(max_new_tokens):
            idx_cond = idx[:, -self.block_size:]
            logits, loss = self(idx_cond)
            logits = logits[:, -1, :]  # (B, C)
            probs = F.softmax(logits, dim=-1)  # (B, C)
            idx_next = torch.multinomial(probs, num_samples=1)  # (B, 1)
            idx = torch.cat((idx, idx_next), dim=1)  # (B, T+1)
        return idx

```

```
[171]: # hyperparameters
batch_size = 32
block_size = 8
max_iters = 5000
eval_interval = 500
lr = 1e-3
device = 'cpu'
```

```

#device = 'cuda'
#device = 'mps' # Mac M1,2,3 only
eval_iters = 200
n_embd = 32
num_heads = 4
n_layer = 3
dropout = 0.2
model = TransformerDecoder(text.vocab_size, n_layer, n_embd, num_heads, block_size, dropout, device)
m = model.to(device)
print(sum(p.numel() for p in m.parameters())/1e6, 'M parameters')

```

0.044059 M parameters

[172]: loss = train_steps(model, text, max_iters, batch_size, block_size, lr, eval_interval, eval_iters, device)

```

step 0: train loss 4.5279, val loss 4.5276
step 500: train loss 2.1601, val loss 2.2149
step 1000: train loss 1.9670, val loss 2.0321
step 1500: train loss 1.8882, val loss 1.9583
step 2000: train loss 1.8225, val loss 1.9032
step 2500: train loss 1.7911, val loss 1.8883
step 3000: train loss 1.7685, val loss 1.8495
step 3500: train loss 1.7506, val loss 1.8430
step 4000: train loss 1.7365, val loss 1.8082
step 4500: train loss 1.7092, val loss 1.7784

```

[173]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))

##ort.

**

Art. 11-2

Cateralas gulicège étémuites artenation

**Art. En staises he reagir à lophabièr baires copritions quz les réputatalité d'une ttruix vicien et lëtition mabite entresant se où une choserrtion êtératéemment par de l'afffongeitt des efffalale de tion du voutre ne quer

```
[174]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=300)[0].tolist()))
```

Art. 14*

Les n'égar la plui, le récévaît qui ier est infisein con reprit caclatan joux préfanus.

Apar les civifemernaclarétabil, du contrans dansser.

Pours étermets ispar de à out l'artire pendivaugis sa nison au jopionage que inatée p pucalien édulation inati-2 à chôte une autrde ou obs

1.7 Réseau avec 10 M paramètres

```
[181]: batch_size = 64
block_size = 256
max_iters = 5000
eval_interval = 500
lr = 3e-4
#device = 'cpu'
#device = 'cuda'
device = 'mps' # Mac M1,2,3 only
eval_iters = 200
n_embed = 384
num_heads = 6
n_layer = 6
dropout = 0.2
seed = 2147483647
torch.manual_seed(seed);
text = TextData('civil.md', device)
model = TransformerDecoder(text.vocab_size, n_layer, n_embed, num_heads, block_size, dropout, device)
m = model.to(device)
print(m)
print(sum(p.numel() for p in m.parameters())/1e6, 'M parameters')
```

```
TransformerDecoder(
  (token_embedding_table): Embedding(91, 384)
  (position_embedding_table): Embedding(256, 384)
  (blocks): Sequential(
    (0): Block(
      (sa_heads): MultiHeadAttention(
        (heads): ModuleList(
          (0-5): 6 x MaskedSelfAttentionHead(
            (key): Linear(in_features=384, out_features=64, bias=False)
```

```

        (query): Linear(in_features=384, out_features=64, bias=False)
        (value): Linear(in_features=384, out_features=64, bias=False)
        (dropout): Dropout(p=0.2, inplace=False)
    )
)
(proj): Linear(in_features=384, out_features=384, bias=True)
(dropout): Dropout(p=0.2, inplace=False)
)
(ffd): FeedForward(
    (net): Sequential(
        (0): Linear(in_features=384, out_features=1536, bias=True)
        (1): ReLU()
        (2): Linear(in_features=1536, out_features=384, bias=True)
        (3): Dropout(p=0.2, inplace=False)
    )
)
(ln1): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
(ln2): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
)
(1): Block(
    (sa_heads): MultiHeadAttention(
        (heads): ModuleList(
            (0-5): 6 x MaskedSelfAttentionHead(
                (key): Linear(in_features=384, out_features=64, bias=False)
                (query): Linear(in_features=384, out_features=64, bias=False)
                (value): Linear(in_features=384, out_features=64, bias=False)
                (dropout): Dropout(p=0.2, inplace=False)
            )
        )
        (proj): Linear(in_features=384, out_features=384, bias=True)
        (dropout): Dropout(p=0.2, inplace=False)
    )
)
(ffd): FeedForward(
    (net): Sequential(
        (0): Linear(in_features=384, out_features=1536, bias=True)
        (1): ReLU()
        (2): Linear(in_features=1536, out_features=384, bias=True)
        (3): Dropout(p=0.2, inplace=False)
    )
)
(ln1): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
(ln2): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
)
(2): Block(
    (sa_heads): MultiHeadAttention(
        (heads): ModuleList(
            (0-5): 6 x MaskedSelfAttentionHead(
                (key): Linear(in_features=384, out_features=64, bias=False)

```

```

        (query): Linear(in_features=384, out_features=64, bias=False)
        (value): Linear(in_features=384, out_features=64, bias=False)
        (dropout): Dropout(p=0.2, inplace=False)
    )
)
(proj): Linear(in_features=384, out_features=384, bias=True)
(dropout): Dropout(p=0.2, inplace=False)
)
(ffd): FeedForward(
    (net): Sequential(
        (0): Linear(in_features=384, out_features=1536, bias=True)
        (1): ReLU()
        (2): Linear(in_features=1536, out_features=384, bias=True)
        (3): Dropout(p=0.2, inplace=False)
    )
)
(ln1): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
(ln2): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
)
(3): Block(
    (sa_heads): MultiHeadAttention(
        (heads): ModuleList(
            (0-5): 6 x MaskedSelfAttentionHead(
                (key): Linear(in_features=384, out_features=64, bias=False)
                (query): Linear(in_features=384, out_features=64, bias=False)
                (value): Linear(in_features=384, out_features=64, bias=False)
                (dropout): Dropout(p=0.2, inplace=False)
            )
        )
        (proj): Linear(in_features=384, out_features=384, bias=True)
        (dropout): Dropout(p=0.2, inplace=False)
    )
    (ffd): FeedForward(
        (net): Sequential(
            (0): Linear(in_features=384, out_features=1536, bias=True)
            (1): ReLU()
            (2): Linear(in_features=1536, out_features=384, bias=True)
            (3): Dropout(p=0.2, inplace=False)
        )
    )
)
(ln1): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
(ln2): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
)
(4): Block(
    (sa_heads): MultiHeadAttention(
        (heads): ModuleList(
            (0-5): 6 x MaskedSelfAttentionHead(
                (key): Linear(in_features=384, out_features=64, bias=False)

```

```

        (query): Linear(in_features=384, out_features=64, bias=False)
        (value): Linear(in_features=384, out_features=64, bias=False)
        (dropout): Dropout(p=0.2, inplace=False)
    )
)
(proj): Linear(in_features=384, out_features=384, bias=True)
(dropout): Dropout(p=0.2, inplace=False)
)
(ffd): FeedForward(
    (net): Sequential(
        (0): Linear(in_features=384, out_features=1536, bias=True)
        (1): ReLU()
        (2): Linear(in_features=1536, out_features=384, bias=True)
        (3): Dropout(p=0.2, inplace=False)
    )
)
(ln1): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
(ln2): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
)
(5): Block(
    (sa_heads): MultiHeadAttention(
        (heads): ModuleList(
            (0-5): 6 x MaskedSelfAttentionHead(
                (key): Linear(in_features=384, out_features=64, bias=False)
                (query): Linear(in_features=384, out_features=64, bias=False)
                (value): Linear(in_features=384, out_features=64, bias=False)
                (dropout): Dropout(p=0.2, inplace=False)
            )
        )
        (proj): Linear(in_features=384, out_features=384, bias=True)
        (dropout): Dropout(p=0.2, inplace=False)
    )
    (ffd): FeedForward(
        (net): Sequential(
            (0): Linear(in_features=384, out_features=1536, bias=True)
            (1): ReLU()
            (2): Linear(in_features=1536, out_features=384, bias=True)
            (3): Dropout(p=0.2, inplace=False)
        )
    )
    (ln1): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
    (ln2): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
)
)
(ln_f): LayerNorm((384,), eps=1e-05, elementwise_affine=True)
(lm_head): Linear(in_features=384, out_features=91, bias=True)
)
10.808923 M parameters

```

```
[182]: loss = train_steps(model, text, max_iters, batch_size, block_size, lr,  
    eval_interval, eval_iters, device)
```

```
step 0: train loss 4.6262, val loss 4.6232  
step 500: train loss 1.2844, val loss 1.3702  
step 1000: train loss 0.9156, val loss 1.0425  
step 1500: train loss 0.8038, val loss 0.9694  
step 2000: train loss 0.7261, val loss 0.9410  
step 2500: train loss 0.6663, val loss 0.9293  
step 3000: train loss 0.6173, val loss 0.9381  
step 3500: train loss 0.5746, val loss 0.9461  
step 4000: train loss 0.5224, val loss 0.9404  
step 4500: train loss 0.4823, val loss 0.9708
```

```
[183]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)  
print(text.decode(model.generate(idx, max_new_tokens=5000)[0].tolist()))
```

Paragraphe 1 : Des libéralités d'alinéagation

Art. 832

Le legs serés dans le conjoint en France présence de l'adoption par un partage, recohé à comptent pour les motifs et s'il y a mis de sa fin, sans préjudice des droits du dépôt que l'affet notaire soit dérogé au règle judiciaire.

Art. 347

En cas de différend ou d'appel, il fera aussitance ou que le décès ou comprononcé des ditémoniers.

Art. 970

Nul ne pourra savoir la réconciliation ne réelle que si le directeur ne résultera de contrat :

1° S'il réside d'un capital ou d'emporter un nouveau consentement d'un assistère public, coous avoir à celui des biens donnés ou des obstructions ou mettenues.

Art. 94

La garantie agricole ne peut être prouvée qu'à dation préatite des personnes désignées revenant les tribunaux pour cet enimeuble autres que la créance personnelles nées dans les conditions prévoitives par chacune des parties, et celui qui appartenaient aux parents à la personne chargée du règlement et de l'initiation que l'intité de l'enfant. Cette entièrement financier à la dation légale de l'exercice de l'autorité parentale relative à la publicité ou à

l'effet de l'acte. La fiducie et des biens d'article ne peuvent contester leur utilités à ce jour.

Tout partie de la prestation solidaire quement de l'époux se trouve dans le ration définie à l'article 21-22.

Art. 28-1

**

A la demande de l'introduction due pendique, la durée de la mesure de sa part ou de la demande en justice a, la déclaration dinjonctive de faculté de ses enfants ou dans un délai de quidentifié son identifiance ou son représentant au remboursement de celui qu'il a été fait postérieurement.

Par dérogation aux dispositionctes des articles 31-1 et, dès lors que les paiements incorportables de l'exercice auparage able intégral ou désignés en cette article 124.

Lorsque l'original d'exatit mentionne que le paiement a procédé à l'étranger peut faire dans l'autre collatérale dans le établissement des services membrorales précentés à l'héritier ou à mesure de celui-ci et affectant de son licitation au plication de l'enle intention allidence et dans deux actes formalisés, par le juge pour exercer une désignation de la violation dans la loi patrimoins de celui-ci. Si l'habilitation administrative du audition famile, la pension peut auprès du mineur pour le commerce un cate corre ou de désignérance, porteant attant sur ce désignant.

Art. 1363

La transaction de patitribution près ainsi que par elle jugarde, la preuve, s'il éta condamner au jour de l'aprète ou de l'eau jour de l'ouverture de la succession, aux créanciers personnels, y compris dans la proposition de trouble mar des biens qui ne doit être réglés sans ou conserve jusqu'à ses impossibilités pour le commerce de la propriété des coïndivisions, sans que le droit de retour partiel ne soit conclu à un tiers gratifié ou à contraindre susceompatif de ses biens ou droits en vente lorsque le conjoint survivant forme indivis ou réduite dans un délai mentionné au premier alinéa de l'article L. 374-5 co-ci, la délibération du caractère civil de la gestion d'affaire sauf à publicité pour y nécéder à cette fin.

L'acte de notification énonce les gérants, même constitués des deux époux, noms, places et consentement dont la remise dont il est établie dans le délai prévu au premier alinéa de l'article 223. Il peut aussi même délité sans que préjudice de condamnation aux mettres choisis par l'un des partenaires soit au gratifié d'avoir effectivement contraires aux articles 1792, 1792-9 et 912.

Chapitre II : De l'ouverture de tuterre

Paragraphe 1 : Des conséquences de la succession est présentée pour la durée antérieur de l'existre et de l'avison.

Art. 955

Un tuteur doit accord, à moins qu'il n'existe passé en France d'un an ou d'impossibilité d'une reconnaissance étranger d'un tiers de part sur l'établissement situé en application des articles 199 ou 1992. Dans ce cas, l'impossibilité d'exécuteur tous les opérations d'aveu immobilique de commande partage interdicte en altération de la faculté d'application d'Etat.

*article 196**

Lorsque la convention de parts se fait par une loi époux ou dans un délai de temps et ou connu, l'autre docuration de l'acte mention sociale à laquelle rend effet à la date de ce délai.

Toute juridiction de l'ouverture de la société ainsi définitive de confirmer des articles 187 et 189.

Art. 1873-2-8

A défaut de titre oppositions à la transcription n'associe justifie.

A défaut d'indemnisation courante des éfereurs, le personne chargée dès un motifé de la propriété. Il peut aussi ou déroger par voie rimmilitante d'un des copartageants et partageants pour faire rendre une olle supposant, qu'il soit agricoles, soit qu'il se disnitôt particulier que par le mariage entre conjoint et par le acte extrajudiciaire, quel le cas échéant sous signature privée ayant été crit préalable.

Chacun d'eux est présumé aux pression du code et de levi-vision, le conjoint se

[184] :

```
idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=5000)[0].tolist())
```

Chapitre III : Dispositions générales

Art. 2215

Toute personne à participation de corps ou devra être adressée par voie émandataire.

Art. 249

Le constituant fait en justice les biens dont il est dû autant que la filiation

soit établie sur l'original de l'état civil fera que les parents effeoires nés dans le délai de six ans à compter de l'officier de l'état civil et sans que le requérant dont il pourra tourner la nullité au préjudice des parties futures aux indivis, les parents se faire protéger avant. Il présenter les une convention d'impossibilité de la survie et des ouvrages publiques sur la partie de l'année, y au autre de celle-ci.

Les partenaires successolifs non comparents, si ceux des personnes de sexe et servitudes et de versement, séparément s'ils n'assuraient d'invoquance.

****Art. 738****

Le rèue des formes du incipes s'ils sont grevés, au jour du décès d'un testament par le partage en lui sera tenu d'étranger sur la part de la délivrance, si elles serouvent à leurs fonds, mais agréé par traianes ou par toustres, faite par la loi ; s'il s'engage à que dessigner.

****Art. 904****

Un testament partagé partiel peut se faire par ses usage de leur leurs, objets qu'après un contenant dans la mesure, ou déterminées par l'article 955, lorsque les coderntités du débiteur originaire les propos dû lit en la supération dans les articles 1057.

Titre II : De l'action complément du débiteur et d'une partie de la chose justes

Section 1 : De la masse

****Art. 1492****

Tout prestation litigiale, prévois par la demande et de dissolution, sous réserve de réserver des articles portants à la date qui le constit, se doutes sous la choses vendues ou pos dû ving de affets, les droits réservataires, à pôts définis, au rapport du bail, sera ce- lié dans le terme domis pour la donation de part. Toute moyenne si, au moment d'ariage.

****Art. 1588****

L'intention d'une contrepartie à d'au isport.

Le preuve est tenu pour le têteur d'être tout autre toute.

Il peut néanmoins qu'il suppourera et directer le territoire qui contient séparatif des anions pour sa valeur de leur biens meubles.

****Art. 1588****

Si, à charge qu'il s'agit et d'une chose, ou de finance que l'usufruit conjoint, ou que le surplus soit conformément aux donations et valeurs acquises, dont il exerce la propriété ou même fraude.

Tout serment converti peut emportant exécution corps l'autre des parties, le garé à compter du jour la donation et employés à la part du fonds et peu hypothèque.

Art. 811

Les règles ceux qui ont été de duré de leur caractère au détruction n'acquiervient que par le défaveur en nature de sexe perdre les biens de cette partie.

Art. 815-14

L'offre est langue d'invention considérée comme par la loi.

En cas de désaccord ; cept, il pourra dix ans seulement motivé d'effet sujetis à l'emprunteur à majoriter de la date et il devra être établi est dans le ressoirat d'actes. Lorsque la mention survie du fait de faute à l'ouverture de l'ouvrage de certaines gors murs, soit en la demande ou exposédérieure.

Titre XI : De l'expertisé en général

Art. 194

Le régime matrimonial est l'assigné avoir au sort des conventions malages du divorce par les mandatages pour une partie à lui rent de l'autorité parente ou sur le territoire exercent en moins de biens donnés, sauf avoir nouf déféré si cessenti à la règle du régime matrimonial, ou subrogé tous réserve de l'intérêt actif au profit de la constat. Elle-même, par que l'effet de la procédure.

Chapitre IIII : Du procédure judiciaire

Chapitre II : De l'attribution par confirmatif au contrat avantage

Paragraphe 1 : Du contrat

Chapitre VI : Dispositions communes

Section 1 : Effois prélègues par

Section 1 : Les actions engages s

Art. 1181

Le contractant est ul en plus ou en application de l'article 1171 ne conférant pas axbe de la donation.

Section 1 : De l'achetique de l'enfant

****Art. 1122****

Le legs à défaut par laquelle une personne, ou sauf de son choix, employé à modifier les planches dix années, au contrat où celle-ci est nécessaire, celui qui lui est due.

Il le vendeur sera aclos au propriétaire des meubles quel que soit la tenue du vendeur aurait droit son plant soindignellement et, à condition que celle-ci n'ont pas provoqué.

****Art. 163****

A stipulation contraire tout ou partie demande de la victimation au conjoint, ou toute personne ayant jaisi devu ou honnombre.

Mention cesse faite hors présumée sur un territoire non responsable ou livré par contrat de fiduciaires et rétroagir, dans le prix à compter de la mort.

Le compte de gestion est temmérieure, au défaut du effet de remplir des lois.

****Art. 1318****

Les services fories établissembles de la communauté ou de participation dans les conventions qu'il a reçues, sont parus.

****Art. 11671****

La respectie de l'acte authéié peut être publiée.

Pour cause de demande non avenir lieu portant de tante, elle mandataires judiciaires, à l'extr

```
[185]: idx = torch.zeros((1, 1), dtype=torch.long, device=device)
print(text.decode(model.generate(idx, max_new_tokens=50000)[0].tolist()))
```

En cas de nullité de l'acte des pouvoirs du représenté construmentant sur la possition exécut de l'une obligation qu'il se hébent point à ceux qui lui ont été condamné de toutes les stipulations susceptibles avec l'assistance du ou des parents du droit incorpopre à la proporiété.

On peut statuer des biens préservatps effets même sur l'apport à quel que le voisionnement n'aisfs à exposerta l'ouverture de la tutelle.

****Art. 549****

Le tuteur, avec l'enjoindre autorisation du tuteur des successions échueus en cas d'assumer autres biens compris dans ce dernier cas, le fiduciaire fera seul nong substanti de disce entre derre eux ou de changement important à autre indigne.

****Art. 827****

La personne entest usentiellement pour les biens présents et indivis ou confirmer une personne, sans que son cocord oblige l'un de certaines d'entier personnes.

****Art. 882****

Chacun des indivisaires peut céder à des mouvertures sans lingte, que l'exécution d'une attribution préférante incidencombale aux frais du maître que l'ouverture de la toute aussituation.

Le notaire au titre des associés jusqu'à préser paraît de la société.

La réparation est opposable au rachat de recours éventuellement, sauf spécifier que lorsqu'ils sont injontants et information allouées compatées avec le dérit en assortie de la contenance ou expressément ce droit au représentant à ses obligations, avec délaissant lui les conditions et logenant dans la limite de la condition que ses deux moeurs ne peuvent être faites dans le cas où le cédant resultané selon les articles 1993.

Lorsques ces pensions aliées en Sont faites perpétuellement en la restriction ne les prouvent que le service qui n'aurait pas cadés de vie d'empris eux.

****Art. 936****

Le lettre greffier de la transaction du défunt de résultat dans un délai de trois mois à compter de la date de la donation.

****Art. 886****

L'obligation est réglée par le décès du donateur, dès la convention contraire.

Lorsque la simple renonce à la quotité l'intérêt du titulage de la fille, le défaut appartienne aux héritiers du propriétaire cédés, le défau de celui donner à la part du précédent s'ils ne justifient, et le conjoint qu'à l'encontre des tiers, au jour du décès si elle sont passédés l'issu eau donant dans le départage ou de tous.

****Art. 17-3****

Lorsque la donation adminieu au décès du dépositaire jusque si la demande est exposée, ses biens ou si meubles, pour assujetti à l'officier de ces biens nés sur la loi faits, le représentant alli qui reprèse la capacité lui et exécution l'héritier, acquiert toute ertouve hypotalement l'industrie peut tenu former l'intérêt de la supprhulation du ou des arryons de famille des confiants.

****Art. 191****

Le mineur est requis pour causes de nul contrat ouvertu de disposer au titre de sa chose soumise aux quêtes louées ou la célébration de faute, quelle enque dure l'une des parties ne rivennent à pourvre qu'ils les fonds, autres et aux créanciers des perstants, désigne et pour le patrimoine.

Du compte et de ces derniers nus par la désignation de la loi applicable les présents et règlements.

Lorsque les biens du donateur comportent au seul effet sans où elles assoliences nécessaires ont effet à l'un des époux ensemble et à causes graves ou verbases.

****Art. 1961****

Celui qui une servitude ne peut recevoir les permières nées rang bonnes et dans la même forme.

****Art. 1769****

Le locataire qui l'en propriét a de chose tournée qui est juste arrivé, soit que le propriétaire de conserver le dhoitage par jusqu'à payer, soit, s'il ne soit donné dans la simple énoncés propre, si la perte est on étangée par suite de l'exploitation contre les héritiers, si celle-ci est applicable aux mices extères.

Les conditions de vente ou d'une partie d'un partage lorsque, les frais, trouve par le venage ainsi invisée, soit par acte mariage, et s'ils sont régies par le code de l'acquet entre les litagements de terre.

****Art. 16-7****

Les fonîtres de l'auteur du contrat de famille ont, à la date de leur testament, directement ou une aquête nécessaire, lorsque les personnes vivant ou alliées ; celle-ci est consignée, servant des marchés constant la liquente.

Section 4 : Du patrimoine et de l'essibilité de la ple ou du conjoint qui ne l'est pas généfit.

****Art. 115****

Elle peut être conclue par l'ouverture de la tutelle ou assurentie.

****Art. 1248****

Lorsqu'il représente un côté plus établit ou au contractant manifeste de le tribunal judiciaire, sous la départie de contradicrinal et si l'officier de l'état civil extérieurs ou au seulent les obligations cultuelles, le procureur de la République de l'officier pouô constituter les fois qu'une exécution soientissantes nécessaires à l'égard des enfants mineurs peuvent être réservés leur enfant mineur.

Titre VII : Des obligations respectives à tranger d'avir au commerce de leurs refus.

****Art. 010****

Pour stipulations, et à titrescrites opérations, qui résultextre du fait de la mauvaise à l'usufruit, un droitier jours de durée que les meubles soiennes donnés que le règlement et à ces derniers ou autres pièces justificaires de règlement.

****Art. 669****

La rend gratifiée a une renonciation envisagée pour cause de nullité du projet d'être exigée avant la nomination.

Si l'une et indivision ne peut être dérogée aux dispositions légales, l'autorité qui était en cas de décès partielle par l'un des parties, quant aux effets de l'adoption en partie de l'exercition simple de l'autorité parentale réglée pour éteindre, l'enteni conserve la faculté réservé à ses faits par le surpropreur, à condition qui soient prive de contrats ou celles que la trampète du donateur ou n'a pas établis que le donateur conserve la nationalité française, n'est pas négociables que le renonçais ne pourra être appéjà huiffraite au juge du tribunal judiciaire que pour qualité, ou préférieur de ces exigences juilles et aurondaires dont il a accepté l'intervertient en acquisition d'un acte de naissance le partage de l'acte de reconnaissance.

Chapitre II : De la néception lieu de faculté par proudre par contrat.

****Art. 1245****

Lorsque l'écrit soit laissance de l'objet sur l'acquisitier. La rente de contrat indéfactive privée devant viable de chaque année, ou par justification, que par décent, soit du défaut d'un tiers donneur, par le fonde de conversion ou d'habitation.

En cas d'extraitssement, l'agent des conjoints, avec les époux et cessentie de proations présentées, pourront être naturés attribuées par jour celle-ci.

Les sième et articles 1963 ou 193, ils sont faits de bâtirs à laquel sevant l'héritage.

A la convention de la sitration en marge faite à la majorité de l'enfant du ou des indivisaires dans la limite de celui où les biens de ces futurs sont ou de ventes, ou si les vis eux ou aux règles de ces biens ou des mentions sociales sur la vie cédente ou aux suscendaires requis portant avec des copartages.

Art. 1817-1

Lorsque la perte s'atteint par actes aux parties exercent colectrifs avec réalisablement à la charge de l'expuls ou aux autres fonctions constatées par l'exerciction de l'ouvrage de l'ouvrage furture des travaux, au pril de celui-ci bres au quelle le vendeur fait du seul epéritatif et dans les fformes useloncées pisonnées selon l'occasion de la personne morale.

Art. 1646

Le tout maître d'établissement, général et tout le fait du comproprement gravel.

Ainsi, la formation de cette indemnité le juge de divorce ou de mettre en iguellié au jour où l'article 182, l'appartient aura, l'action passéfaite à l'officier de publication déjà sont exercées à cet expert.

Français conformément aux formalités de publicité foncoire d'une attribution auprès des conditions relatifs à l'acte qui en permet aura été conformément au diuxième alinéa de l'article 401, si ce cet acte est requis par jours de la notification établie par l'Etat, contractant de l'enfant pédce d'un des enfants ou de l'une des parties au notaire s'il n'a qui en pouvait procéder à sa période d'un bien dont il est au jour du juge.

Art. 101-6

Que cent action relative à la filiation l'enfant contractuelle ou à l'intimation, électron, et que toute autre ascendant itellement construite et quel que serait nécessaire, le juge aux finalités et ordonnels de gestion.

Pure et simple, elle directeur des services de fonction et des conventions susvisées par les articles 1079 ou 109.

Cette formalité en accomplit l'extissement de ceptation de fonction que l'octroie de la personne ou le nom aura transmissible de la rupture du mariage déjà courrait pour à ses regardes, de l'établissements judiciaires, dès qu'sur les cohéritiers exercés en l'absence des deurs factions par l'ordre de contracter pour inscription des bârites : toutes parties, ou par suite d'ouvrage et d'autres étravatux, s'en trouve que l'âge fait ou du mineur entier dans un bien.

A défaut d'entrete de l'une des parties, le présent contrat est requiertant l'établissement du partage et simple le compte cessionnay et licite en application de l'article 279.

Dans ce cas, le légataire déterminé par un comminique de plus procéder, la date et le lieu du dépôt donneu en appartenant intère dont il est défini aux articles 1973.

Lorsque la clôture n'est établie par arrêté dans le délai de six mois.

****Art. 1153****

Le testament par lequel ne veut aucune puisse être formée, de l'ouverture du testament par le titre l'autorale n'ordre.

****Art. 1983****

Le legs sera prononcé obligé dès la cause que la nullité du seul six mois, au régime matrimonial.

Il est présidé au tribunal du constituant le fiduciaire peut égal constituer sur l'établissement requis le rejotant du défunt d'ouvent non acceptation à condition qu'ils ne soit nécessaires.

****Art. 71****

Toute personne bilitée peut dons plus jugées en le partage doit résultant d'une destinée à partir d'antourgenter les personnes qu'ils constantent ou réalisés dans la gestion.

Titre XI : De l'extinction de l'indemnité des biens

****Art. 515****

Sont différents à un établissement des associations sociapendant chacun d'exiger qu'il a recevoit, à moins qu'ils entiennent, après que le conseembre de fonds partenaires, ou plus dans la limite d'un décès en ligne différant ou celles-ci aussi circonstanciées.

****Art. 1322****

L'interpellation peut être faite à l'officier de l'administration légale ne sera plus recative qui lui conforme à l'entreprise.

****Art. 1320****

Dans le contrat est sous réserve de plein droit inopnible.

Celui dont l'authe.

****Art. 1101-3****

Est la nationalité française, sans déterminable le décès et justes pour indice, pour quelque que ce soit, les travaux ont nécessaires sont pareu autorisées sur effet de leur communt, par le vice cessoire, et par licitation de la source de fonction.

Il doit être porté dans les cas pigences villes sont impossibles.

Lorsqu'une mention d'incorporation entre des faits partiels sont alliés par donations particulières, les personnes électes pendant l'ouverture de l'établissement d'une commune du commun acte et de geroresponsable du fong profit dans les trois conditions premiers prévues au premie alinéa de l'article 375-11.

Toutefois, le directeur des servitudes exeettenties pour fachager les cations passées peut être dressées hors dans le repect de la situation.

Les faits par la demande de visite pour toute rente viagère ou son autre dessaisin exposés sont sur sa faute et sous sinaire doit de faire visigner les mineurs à 900 qu'ils reflets.

****Art. 9****

Celui qui a reçus dans l'héritage par l'héritier avatage qu'il n'a pris jour. Dans ce cas, le cas échéapeur venant la sommation de l'obligation, les propriétaires s'engagent pour l'objet ou la réparation du donataire.

****Art. 1047****

Le donataire jouit à la chose qui l'usent, et pour le mur tant d'usage par luit.

S'il définit au cours d'eau domage pour l'identification de l'usufruit concubine.

****Art. 577****

Tout propriétaire produit ou soin appartienne entre le propriétaire et l'autre, ou le dol, juge motifs, que doit le tribunal justifient propriétaire, après avoir été contre eux assemblé de son décès dans les conditions prévues à l'article 983.

Cet article peut demander à la présomption de rémunter en application du premier alinéa de l'article 418-1 du présent article simple à la date de la mesure si, qu'ils soient établis au jour de la vérification s'il a sont encore publiques dès lors que ceux dont ils sont déchargés par décret en Conseil d'Etat. Tout

assemblé dans ces formalités. Il peut, à défaut, et, à un changement donné de corps, de placement exception ou scelle dont il a été versé par un auteur.

Titre IV : Des obligations

Chapitre Ier : Dispositions générées.

Chapitre Ier : Des actes de déses requisentre le national et de la découverte de la part de louage dans un lot du marché ou des militayés dans l'article 74 qui en ont été renonçant négociés.

Art. 756-5

Toute clause des disposants appart en nature en application de l'article 1794-3 sont entierdants à l'égard des titres de pereurs, soit mention de leur acte divrant notas.

Art. 229

Au renonçant, le président du tribunal, la restitution du entretie contrain d'un des doctions exerçant la fonction de la mention.

A décision justifiant de résiliation prévue à l'article 229 et que le partage aura été constaté servi cette décret n'est pas sans part.

Art. 223-2

En franche d'une délégation prévue aux dispositions de l'article précédent, le juge peut décider que l'exécution du conseil de famille sus.

Il suffiera de mention d'auteur pour monter elle terme.

Art. 333

Le directeur des services de greffieraiement de préeférenté ou consentir au propriétaire livre de bonne fonds inférienteront.

Toutee statut légitime des personnes détachées à l'ordre public pour les effets de la conficitude de la tutelle.

Art. 342-2

Par des unes époux lorsqu'il est indivisale à l'un des époux dans l'indivisaire .

Art. 832-1

Un indivisaire sera résidu par l'autorité purement commun présent les

circonstances d'un deux partage ou indisponible.

****Art. 813-7****

L'indivisaire par que le médecineur d'avoir pris de gré son parts exu et portels, plus déjà, les faits ne pourront être contestés sur les biens personnels avant l'ordre héréditaire ; à accas de discru présention d'un mandant non définis à compter de leur proposition et des attants définissements de la personne protégée ainsi qu'il aupara donner cette disposition.

****Art. 107****

Quelque qualité de manière dont les parties sont respectives lors ou qu'ils seront rapportables les actions utiles qui font la part.

Titre IV : De l'obligation par lettrantier.

****Art. 1761****

Celui qui qui restera salour propriétaire entièrement l'obligation porte une moindre de la succession entre contre les matières.

****Art. 1746****

Celui qui veul prene nature est loyeur dans les ventes, une baage soit doute autre ; et ils se faient à l'ouvrage, s'ils sera parlés ou l'exclusion de juge.

Ainsi, le juge des tutelles tutélaires sans que l'on claisse procéder, due à la date ou aux admettres.

Si le remplace apits de finance ou de gérance, les légitimes du contrat s'il procèdent d'une prestation générale, sera faite avec les époux à peine déterminant dominant des parties dans la partie ou sous la convenance que si elles sont contraires. Le juge dispositif est applicaté par le contenu avec tradition lorsque la libéralité foncé, la valaudie des sommes esse constitue entièrement.

****Art. 116-4****

Toute nouvelle dénonciation, exécuteur testament olidant son usage attual défini d'une nationalité, par le changement du régime matrimonial, que si celui-ci py font la personne. Il le servit sera à liquidation des dispositions de l'article 912 du présent code. Il défermet déterminant l'option qui le dépôt n'est responsable au décès auraitifié le droit de grossi patrimonial.

****Art. 978****

Si le titre conominat la chose légot indivitée le dépôt jusqu'à ce que les conditions dont le dépôt nataire ne s'en pas prête, et qu'il n'en produit

qu'autres autants que le prix, et du jour de ce qu'il est exigible.

Il est par jours par le prix pour le rende du déposant.

Paragraphe 2 pour compter toute la réserve.

Art. 958

L'erreur tout ou exation pour sa mort, l'article 964 peut être protégé) totarde en tout ou partie, le donateur ou de l'établissement d'utilité publics ne peut retenir le capital que le donneur en lieu cas de meublération ou d'après la conflition publique anticitée.

Section 1 : De la distinguition de l'objet ou des parties.

Art. 1832-1

L'intéressé peut de faire contracter mariage sans le consentement de chacune d'entre elle s'en absence d'une part.

Le parent peut révoquer la novation qu'"elle décède avant qu'elle était entièr au moment où celui au des enfants confiés au plus de rente. A la demande de l'un des parties ou de l'autre parties lui s'agi, le juge peut prendre une délibération suivante contre les parties.

A défaut de stipulations du contrat alors avec le consentement au règlement de la prestation compensatoire, à la communauté, ou à la suite du logement. L'écrit électronique à celui dont l'exploitation des dommages l'intérêt possible et s'il n'était qu'il lui a jouissance.

Lorsque la demande en réduction est, le droit de portion soide mention de l'équipement ;

elle est subordonnée à un préjudice carté à par le créancier personnel de la protection de l'existence ou d'un délai de trois mois ou soins après la publication du procureur de la République. S'il est de faire mation sera en cas de nullité.

Art. 574

Le juge de séparation de biens alors que la liquidation de la mise privée participation ou certainelle entre les héritiers usu habilités à payer les créanciers de sa part. Tout ou des biens particuliers indivis, également à l'âge de décision tutile et d'au-depuis six ans.

***Art. 789**

Les mêmes différents ou services ont tenus arrêtées dans un pays autre plus ont tenus :

S'il l'a état de son première ;

1° Les conséquences de la personne physique et le capital sont considérés comme partages.

Art. 982

Les héritiers peuvent exercer les pièces et immeubles.

Lorsque les dettes, même sont rédictés ou ciractés civiles, aux actes faits en lieu entre, la note, l'exécution des conventions statutant retifs ou de trois ceux ; et celles peronnes des bêtes, par les sommes baries ou intérêtes.

Art. 191

Lorsque la chose umitoyenne a fait seulement résidéficiée au tribunal, par l'effet du la vice d'un parent ou sa forme nature, les parties sont acceptées. La devante est ou la réduction de l'originaux tiers de la subrogation, pour l'une et l'acte de partage, la chose est est contraire.

Art. 1932

Toute les parties par cas fortuit de la terre en part suffisante à celui qui use pour se prélève en cas de nullité, le juge du conjoint sera défini par le caractaion.

Chapitre IV : Le dépôt prévoqué dans les conditions de la loi

Section 1 : L'exception " Dispositions contestantes lorsque l'altération de facultés d'actes ou d'approbation insuffisante d'un délai de six mois, pour l'une des aliments, faite et l'objet une clause d'un vocation pronoée en sociéléère dans une uniltération officielle de l'état civil qui est opposable à l'instance de deux aïeuls, l'acte est réservé de l'officier de l'état civil du dernier droit de dernier comminé lorsqu'il n'est pas décédé ou sur les trois relatifs au complément de la date de ponsibilité forcée.

A défaut de convention, la pension en sera faite par l'un des futurs époux et le président du tembre du " R° du contrat et, avec la mission ou la mission du corps ou s'il devient à ses droits d'ouverture tants, et ils jourgent, sur le formie de la publication prévues au de l'article 92.

Si elle a expressémenté est expliquée par l'acte est sujet à primite du divorce restournt pris partiel.

Si le recouvreur n'a dressé toute pérition d'une cours de la paoption provisoitive.

Chapitre VIII : Du contrat et d'adoption résulte pour une durée indénationnée exce de la prescription ou par l'aliénation ou sur les registres de l'état civil du donateur offre qu'il sparaissait nécessitané de leur faculté de perder l'acte cérébré et les têts, droits et actions sur les Françaires et sœurs au ministre de la célébroation du mariage subsistestés par acte publics ou par authentique.

Art. 191

Ces peuvent être droites judiciaires, ou avec décret souscrité pure les assignations sur un titre ques n'on oblige d'acheter une mention de défenseil du territoire fixé es permettant dans l'intérêt de la défaculté de reconversion ;

2° Les exerûtes gruement rés français entre productions ;

2° Les planchandiciens du défunt saninces pour une faute incorporant adoption simple n'excède pas suffisante.

Art. 931-2

A l'étranger, dès le retur de la naissance des décents d'autres père, modifié ou sur raison de son chamin et de plein droit.

Art. 770

Le dépositaire de la personne physique n'est pas interdit par le pacte civil de solidarité, et qu'il fixera aussisté plus de plein droit de réserve pour l'actif sur l'autorité publique.

Art. 830-2

Le pacte civil de solidarité prend fin à l'officier de l'état civil qui n'apparait que la famille possible pour faute amiable acceptant.

Avoir enfant le testament mystique à ses intéressés, le trouvement cente menation.

Section 3 : Des opérations contractues par les articles 1180 à 1186.

Section 1 : De la cession de corps par acte extrajudiciaire, ou le fiduciaire fait procation sur un registre que les témoins qui ont participé à l'un ou à titre ulaquelle le consentement ensemble de l'instrument.

Art. 1901

Le fiduciaire peut être contaté ou déterminé par le contrat de conseil de famille, toute audition que l'enfant ou l'exercice de l'autorité parentale pour

le mineur des filiats en vue de sa majorité ou en différence l'intentioncation, activé est également relative à l'établissement de l'enfants et à cause du divorce.

****Art. 36****

Avant l'officier de l'état civil du devien de son adoption si conformés pour les enfants du conjoint et auteurs demandeurs ou aux dissiuués par l'article 311 conclus.

Celui chacun des partenaires peut, pour la saisine de souvrait à celle de ses biens expliqués, physiques ou marques, selon les règles de partages participarticipales, des dispositions liblitaires.

Titre IV : De l'établissement de la filiation de l'emporteur

****Art. 712****

En cas de retardé ou, l'office expendant un profit génétique.

Les servitudes spicabiles est réputées non équivoques sur ceux-ci pour leurs deourges gérants, d'être tenus par le fait de celui-ci, une servitude de paossier dans l'héritage de ses biens ou des droits dont le fonds acquittaire la valeur des biens immobilistrats ou à usage qui soient occulés à un crédit ou un certain détacte n'existence que le jugal ses diffifications. Il peut également autoriser le projet du depuble du conjoint étranger.

Il doit aussi jête privé, ingfé de la limite, qu'ils éclarent ou non saisissance, dans les conditions prévues par le présentecteur.

Les actes passés en contestation dès que l'annul puisse ou la décision de six mois après l'expiration de ces dispositions.

A défaut de déclaration carant, la date ofertunale s'est rendue postéjà sous indivisible, en son champ.

Constitue d'une action nonGidentement li réclame la personne à protéger le subrogé tuteur et apprés au juge, jusqu'à des héritiers naturels ou qu'il a réalisé la communation de comptes ou de dettes.

Il résilite à l'égation des immeubles gérants et à publiciter ou à la fin du des parents, celui qui est impossible par son représentant légal eudes arrêotes.

****Art. 914****

Le mineur âgé de meubles ou immobiliers univeaux sales et besoins gales, ou pièces, comme unezsses poids, crolières ou taches, tous fruits exclus des rurs successifs, et publiés tant dans un terme forge, au titre " Aitre " Des

servitudes sont passifs ou françaises perdés.

Lorsque les époux sont exécutés en confiant de la signation ou du domicile légué dans un pays apporte le régime matrimonial entière de l'autre époux et par le ministère conjugal.

****Art. 109****

La mession a été procédée par voie déemandant de conserver sa provision et de répondre d'un autre, ainsi que le jugement d'aura, en présemplice des déclarations sont requises.

****Art. 91****

Lorsque deux époux constituant une demande en juge lorsque l'indivision d'un usufruit a été prise placée, la constitution ne peuvent préjudice que sur une faute de l'objet.

L'usufruitier ne s'engagera pas assisté de réductions devant du terrain qu'elle a pu expiration avant le tems de l'usufruit, ni degr s'en sera exercé nullu), et les droits dont la succession, les héritiers qui n'ont pas pris les luits soins ou n'en concerne qu'autant que les premières d'assituation de cette produite.

Le production de cette ordre est fixé à l'article 648.

Chapitre XII : Des successions et de la procédure de mariage, soits, ou s'il n'y a pris héritiers.

****Art. 57****

Ne peuvent être constitués de ses lois par les optiers, ou par l'effet d'un tiers ayant payant génér d'une autre.

****Art. 792****

Lorsque terme nature qui s'affecte dans les conditions de l'article 816, chacun de ses époux, et par contribuer aux pouvoirs un partage.

****Art. 811-18****

Lorsque le propriétaire fait procéder à la mort des autres voirs de perte ou de faire, l'attribution de l'opération ou des choisirs dans la limite d'une obligation professionnelle et de fonds dont ils y sont constitués par les uservices recelont celles des pareilles emportent les articles 1792- prescrites arrictices alors.

Chapitre II : Des condamner à établissements de police ou consistant dans la gestion relativement au conjoint survivant ou tant qu'ils demandent séparés

par un partage sufisant pour réclamer le préjudice pour le cas d'usage qu'il affecte.

****Art. 1201****

L'impossible de plein droit nature la capacité de recevoir ou contre le droit du jour où il est opposable, pendant à la transcription, à la réalité, par les autres articuliers, l'acte de l'état civil est membre des formalités de l'exercice de chaque partie les droits d'un Jurnat à la commune où la date et le lieu est ordonnée.

Chapitre II : Des conventions entre les gains qu'ils profitent et s'incombeablement.

****Art. 1049****

Dans tous les cas où la commune où un époux ne souhaite placé dans le bail.

Toutefois, l'époux viage qui se enti au jour de la communauté, on donne reconnaissance à celui ou à son propre, dans une forme ou autrement dont il n'est pas tenu pour la présomption tout tout ou partie en application de l'article 1438, sur la garde prévue par l'article 1388 et aux articles 1873-1 ; lequel acceptation en cas de réparation.

****Art. 1873-4****

Un héritier valant caire sans optable le partage, ou celui qui a exclu un domandé, sans écrit, y contre celui dont les titres a engrevance de leur rente viagère. Dans ce cas, l'héritier ou y conditionnellement a passé force. Il en assure électue à la responsabilité des indivisaires.

Le débiteur dès tout moment de l'introduction patrimoniale dont le défunt est déficiable à la destination opérée sur l'objet vient à la mitoyenneté. Elle peut également devient faire à la personne chargée.

****Art. 798****

Les sépoux pourront se point. Ils cessent de droit en national, dès que la forme anticipée, au lieu et héritier, à ceux qui n'ont pas indivisionné leur copiennel indéfiniques, héritiers ou risseurs désignés par la loi en la forme de la transmission des biens et aménés ou des personnes ses.

****Art. 919****

Les droits sont de santé et dans les conditions fixées par l'article 91 eu énonce les intéressants et frauds dissous.

Ces actes de l'état civil est assuré dans la résidence de l'acte de naissance de

l'enfant. En cas de mesure de vie ou de caution immobilière espéciale pour cas d'application de l'article 311-2 alinéa), le consentement à l'adhésion de la procédente sont applicables à l'endroit de l'indignité mentionnés avant l'adoption prévue à l'article 319, si est mineur à qui l'adoption est âgée de mettre effectué pour les actes d'adoption on coreflectives à la personne empreinte, par représenter du défacord ou l'adoption de siège lui de l'adoptant au jour où ils sont optés.

Titre X : De l'adoption n'est établie qu'avec la forme d'un caractère de détermination entre le titre de l'autre mandataire.

Art. 368

La personne charge du droit de commercial ou de l'adopté entre les nationalités dans les conditions prévues au premier alité de l'article 342, sous réserve.

Toutefois, si la satutâ cède est indéterminée, nécessaire, le cas échéant, le régi et à la dernière délégué aux noms de l'enfant.

En cas d'adoption par des enfants âgés de modification du conseil à exercer l'ouverture de l'autre des tiers ; sans que ce ne se soit fait en cas d'usage retournement à l'époux qu'il soit de manifester suspension de lui personnes.

Section 2 : Des mensommes qui saissent sont

Section 1 : De l'annuant du jour où

Art. 1999

Les lots, ess opérations sont encours applicables aux constitutions publiques, sont déposées contractement par acte notariés par le juge.

Art. 1917

Même est notammé, sauf condition aux règles piles fixées au chapitre II du présent article.

La renonciation prévue au 2° de l'article 1126 et sur résidence habituelle à la date de la rente assistance au maintien de la personne décédée.

Art. 1402

Le mariage de la contribution par consentement mutuellement est prescriptible avec solidaire.

Art. 1327

Lorsque le désisposant est nécessaire pour le remplacer, la répartition d'aborder

à sa caution relative à la remettre n'autorise ; la suparaisse de donner la dette et achève du cédant.

Lorsqu'il y a de demande, payer au alluvement de demande en nature en fonction des dépositions graduels, et répondant indivis, à moins que le testateur ait constitué sont en outicvement, pour l'usage de l'acquéreur, le fermier ou la découverte des négociations.

Il doit défender que la somme éconstataire soit hors sur l'option sur ce moyen tableaux, le rachant de payer au-delà du partage, s'il y conserve tout ou ses sindes conformancés par des lois militaires, sauf à l'offre et à l'égard de son conjoint et qui doit rester l'usufruit.

Section 5 : Des deux propres dont le père se fait en surplé par le fait et le mariage d'exploitation.

Art. 951

Les créanciers successorals sont applicables à l'égard des lieux, à l'exercice de l'autorité parentale relativent à l'enfant ayant fait.

La demande d'attribution effectuée par article conseil agréé réservée portant personnel l'un des parents associés ou d'un associé continue de corporer conjointement au gratifié ou indignaturel des services confiées.

L'associé éte valance de ses parts socia pour sa gestion anticipée les règles privées à l'article 471.

Art. 811-11

Les dégradations des parts sociales sont applicables au conjoint sans où ces garants compris dans les avoir exercent leur portants au dégradation.

Art. 181

Les jours de la nationalité de la sointé.

Sont admises possibles d'un acte sous seing privé par l'officier de l'état civil français, dans les deux alinéas précédents, Pour les associés par eux en application de l'article 135, la convention prévue au chapitre VI du présent titre sont applicables à cette derminer.

Le mineur notarmé que produit de ce dernier alnnées n'est pas effectivement pour rapporter le profit.

Art. 1841-9

Le juge peut, sans préjudice de dommage éciable, sur son conjoint, par voie de

cette commune association, lui et recours cette mesure.

Aucune héritiers dans un apport ou force matériellement, si celle-ci sont aggravées dans les conditions fixées par décret en Conseil d'Etat. Dans le cas prévu à l'article 163, le domicile dans le délai de deux ans à partir des actes de dernière ou des personnes ou de subdister ayant solidarités six celles du père et tant qu'à courir la et de l'exercice de lequel l'action est constatée en constituant titulaire sur un délai de tribunal.

Art. 1367

Tout personne me est tenue à agration de sa famille, si l'une de ce de signature, sa des fruits revenus coups ou s'il n'objet pour une consomme à usage, le juge poursuivre le deur, nom d'enseignement lieu conclus avec à l'étranger pour les tribunaux affectant l'origine du lieu où le ministère L. 197 et 271 profit de l'autre autre brangør peuvent être compétents annuels.

Chacunement des baux appartient à la porte sur hacune et qui s'appraité à préjudice écoulé ou commodetrait périt, il industre quelle que s'ils les tiens :

Les autres parents s'exercent sont à l'égard de celles-ci.

En cas de séparation de corps ;

2° parmi le droit de faire à un tiers, à un autre libéral ou au seindu, à due fonds ou un autre spèce gratuit.

On ne puisse saver.

Le consentementement de tout ou partie sa décord.

Le garanti est obligé avec lui, vendeur usuffisante, ou que si celui un artisanan avaité ou acquis, que peut y déroger aux pièmes et circonstances, le juge peut si, par cedtion ou stipulares, soit par acte de titres ou parties sur la personne, sous ressorts de l'encontion à son domicile et aux observations des dommages et intégrales.

Chacun d'eux pour les pouvoirs exercent sous réserve des circonstances propres partionnels.

Art. 810

Tout legs fait él'ordre des articles précédententuels, sauf différente, obstantiellement urgence, tels ouvrages, telles glans que l'on aurait fait l'objet d'une clause d'héberge ne montané pas la plus réclament en cause.

L'opposition poseère, en cas de mandat, notamment conclu par certains sont spécants au jour au transmettre ont directement cédés :

1° Celui duquement effectué de séparation de biens alors ;

3° Exercer ou de choses héritiers.

Section 2 : La légale du mandat successoral est défini de la nationalité française pendant une moindre séparation ou d'émancipation supérie de l'article 225 et réalisée au degré ins du mineur ne peut ;

- dans les cadres excès de la mention ou de l'acte juridique exécutés au fichier immorité est spécialement réputée par exe, celui du invenaire qui a été par preuvenir précodé. Cette signification est prescrite par le code de procédure civile.

Art. 2244

Lorsqu'un immeuble de plein droit nantéa de sa majorité si la loi est fondé ou suivant l'original de la transcription.

Art. 217

Lorsque la communauté istuée est hors de gérer, le partage s'il élègue et si le conjoint survivant son époux, faire à titre d'acte gratuit ou lorsque la violence à l'épouhére sur un bien indivis au profit d'un indivisaire dans sa demande.

Art. 1248-3

Tout mariage présentatif à l'article précédent s'il a obtenu, elle est mineure exigée contre les père et mère.

Lorsqu'il est fait le desquels formes de l'intérêt qui résultera de l'intégration engagard le titre visé du mariage.

Si l'héritier n'a pas été contraste pour un acte de notoriété sur lesdites de profits requières le conjoint se trouve hypothécaire, s'il y a eu, entre ses facultés de celui ou celés de leurs défunts de l'assistance médicationnées ou son reconnu ou à cause de loguer d'engardement dans les conditions prévues après l'article 10-8 et aux donations prévues au chapitre, nonobstantement toute déclaration aux donataires du on autre, s'ils conneront à ce qu'il est connu selon et à liquidations.

Un dol litige et dans le forme d'une partie telle, conversible, à rachat pour contre celle qu'elle se prévie.

Art. 897

Le preneur d'une rente viagère ou jusqu'à concurrence de saisine déterminable

avec lui de situation d'usufruit telle : ce lot, entre les autres, cohéritiers, arbants, la sanction ne peuvent toujours s'appliquer que leurs propriétaires.

Chapitre II : Du versement des capitaux sociaux que profitée peuvent être contesté par ventesté soit due contrad acte unilatérial.

****Art. 1873****

Le montant ne peut être révoqué que sans qu'il y ait.

****Art. 1847****

Le principe contestation peut être mortée contre les tiers de conformité de participation.

****Art. 1893****

Dans ce cas de séparation entre les époux, lorsque le logement fait courir la dation lui pour garaches par celles ou collectifs d'un fonctionnaire.

****Art. 1166****

La contribution est de portion relative aux effets d'une partie d'ont réunir aux quarchier ce jugé estimatif ou transfert de donnes fois le preneur vivenu ou une cationne.

Lorsque la vente puisse prendre la même forme une soulte.

Ces frais exclusivement est tenu de rendre tous les efreveurs et incédés s'ils y appartiennent.

****Art. 1681****

L'héritage cédé ficie qui ne peut être renvelé qu'en suffisant postérieurement au procès du venaire de l'objet de prendre une chose pareille à un fonds économie instant et porté à son difice, mais, en la fonction que celle-ci assera le fait dans l'impossibilité d'un autre compate et non entièrement.

****Art. 916****

Lorsque l'impossibilité de deux mineurs ou de faire chaque héritiers est instruit, un fonds de commercainoire, exclus table dans la proportion aux cours clos et immeubles d'utile. Elle est nul le lieu à procéder, quelle que soit le juge compétent.

****Art. 1821****

Les gains ne soit pas sociés chez unes sociées, soit qu'ils ne peuvent pas

prononcées. Cette agrication énue s'associent passisté une expédition judiciaire. Elle s'impose aux héritiers ayant condamnés pour cette sommation conventionnelle pour la cas durable, que le logement aura conventionu, l'amendeur ne peut être fait que sur l'ascendant de celui d'autres personnes.

****Art. 1926****

Le mariage peut seulement déclaré être postérieur à la créance majorité, par le défunt comprenant sur réserve que la délivrance de l'oncle et est insuspendue de ses confientements au propriétaire, avec qu'elle-mén, par le cassoral ou des terres vifs ou plus grevétiques.

Il comprend les indivisions qu'ellent, la loi applicable comptante est témois de la partage et dant indiquant de l'Etat, à la responsabilité de la régimité, au fiduciaire d'au moins un Et amétique à la disposition ou à titre gratuit civeI.

****Art. 1070****

La personne à bénéficiaire s'opère à s'aux coproportions que celles-ci eûnt libéralité soient retionnellie par une acte conjointement : de l'effet de parié consentement à l'adoption, trois et, parts dont le conféré successionne que l'ordre de l'autre parent.

L'alliéation ne s'exerce pendant anticipens aux fins de suite mauvais.

Sont cessés ptoires sont présumés obstaclés avant leur ager.

****Art. 39****

Les droits du futur époux compromettent les autres d'entretien.

Ils sont dispariés par sonne entitroduier. Toutefois, les copropriétaires des tiers étournent propriétaire de tout ce plus arrière, pendant la communauté un pré de leur époux sont, eux en cas, soit à nul excéder un moins, lorsque les formalités et placés de possession sans solidarisation sont mais, en cas d'acquitter tout créance lorsqu'ils constant d'un cohéritier.

****Art. 813****

Le rachat par le créancier à concurrence de celui qui vient exigible, le juge peut rendre un retroitoir, si elle est tenue au point d'indiquer pour donner le prix.

Si le bâtiment professionnel titulaire en l'absence par le créancier, s'il en exprime la part, il les a faits de la capart.

Il n'a pas d'avance faite à l'égranche.

La novation ou l'opère soit résouter la condition à faute de l'autre quelque par le notaire l'offre punitie des tiers.

****Art. 579****

Dans le chacune de certains son acte ou de force excéder également fondé l'ordre altÉ tat adjuger aux fins d'une action successival.

****Art. 756****

L'opération héréditagée peut être forhécialement par lui de son cochoix dans la limite exprimée, sous le cas fortuit. Elle est payable par le juge à partir du jour de la famille, de la location ou du juge.

Chapitre II : Du contentement du dépôt et du demandeur si en faveur de l'ordre mobiunaux

****Art. 518****

Toute personne majeure a été d'audition de l'ouvrage s'engagé en cas d'impossibilité procédant par voie de l'annulation du devant ce commun.

****Art. 1211****

On peut attribuer à la saisine d'une fiduciaire adjudicte apparaès l'ouverture de la succession, si le légataire n'est pas plus inconnu des nouveaux alinéas et applicables ; eu des frais de l'article bénificien de la propriété ou d'une personne met. Elle s'oppose sur l'étranger, elle nécessaire à tout acte ou à dettres fins qui comporteraains l'assistance ou la possibilité de l'exercice de l'autorité parentale.

****Art. 98 tefois pour couvrir le titre mentionne moindre les fins de sa reconnaissance patrimoine, ainsi qu'aux peines opérées chomaises et totalements de police apsychacun du chef ou le représenté, qu'il y a conservé à l'assistance période, déterminée, ou du jour du décès, tant au représenté d'exécution que la sera rendre.**

****Art. 96****

Le présomption à titre ou fait ne pourra, ni être tenu sur déterminé de ce dont eltement, il doit être dû établi que par le donateur sans ressort.

****Art. 959****

Ne poursuite possession du contrat n'est plus applicable au créancier les autres cohérichtions et ayant cours avantat les cohéritiers de celui ou de son ayants que la participation de l'acquéreur dant le donateur n'exerce que par sa dette,

ils peuvent, offrir au représenté de leur établir. Toutefois les formalités légieuses sus dans la réparation du rapport ou entre une somme édictée faite sans part que l'acquéreur, l'acte qui a risqué un engagement autrement d'événements conservatoire.

****Art. 1693****

Dans le cas de lui donation ou de fermier ou d'effets mobiliers à garant où l'usage aurait la charte, et s'il n'a pas moins de six notaires en remboursant consentir à la recueillir ce qui a réduit.

****Art. 977****

Présence des dispositions, des articles 987 ou 76, matrimoniales négligées à l'article 97-4 ou, les époux peuvent convenir que le jugement permet fait de l'introduction d'affecter la demande sa visite adressée, en jugee des enfants doit être regardé dans les conditions prévues aux articles 276 et suivants.

Chapitre II : Des dispositions générales exprimées dans la date à sa portion, entrant la demande d'un cohéritier, doit les exceptions jugement, soit d'office, d'état souhaitement également, soit en faisant appartie une contenance du prix aidenu de cinq années de filiation justifiée au fichier de l'ordre célèbre expert ou où ils doivent indispositifs de l'indemnité.

****Art. 1793-2-1****

Par dérogation aux articles 1901 à 1699 qui appartient de partie ou de la libéralité de leur partage et dont le consentement était en apportant au service décédé dans des circonstances évitables et chargés à cette règle intervention des articles personnes à qui a renouvelé un assemblé dans les conditions du présent chapitre ordinaire et pour la gestion des successions.

A défaut d'administration portantale, les pertes en engagent le droit retardiairement d'homologation sur le débiteur de la situation principale, chacun d'eux donne au plus réglé par l'article 902.

A défaut de juridiction contraire du tribunal judiciaire que résilie de l'autorité parentale conserve l'exercice de l'autorité parentale.

A défaut d'une parentale mentionnée au sexe ° de l'article 304-2 si, le conflit est réputé noncé de filiation à l'égard d'un collectif d'enfant ou d'autres âges. L'option qui doit être remprochée par écrit ou son consentement expédition de l'article 1370-3.

****Art. 373-1****

La filiation prévue en article 130, est supérieure à la filiation pendant le montant et que le défunt est partenu au pareau, avec le règlement de l'intéressé

diffsaire ne sait pas représenté, n'ont pu plus.

Titre VII : Des meubles.

Art. 340

Les servitus, dans les contributions de l'entreprise si le propriétaire lorsqu'ils ayant pour objet contre défidération, lequelle ces règles prescrites pour les obligations du contentible.

S'il y a été faite de biens parties, non vendu caditement sur le territoire français.

Art. 971

Dans l'entotaire mystéral nationale, l'importance ou le juge qui statue sur la liste de leur valance des objets qu'il prouve.

L'acquéreur dans ces biens communs profits avant ou modifier le diffirant, quel que le vendeur ait eu l'exercice de ce délai ou particulier, de la rélèvement corps, réciproque des copergants ascendants à celles de qui contrat.

A défaut d'héritier les deux lotaires et justifient des instants mentionnés aux articles exprimées par le cople-là préférence dot, mais nonon-usage compromies des immeubles à un réparation des biens qui aient été déduisous, au jour des objets compris effectant maie des liers de l'ouvreure d'entretien.

Lorsqu'il s'agit des mentions prévues aux articles 167, 16-1 et 27-5, la modification de présidence prévue au 2° de l'article 516.

Art. 21-15

Le projet de l'autorité parentale exer à sa présentation selon les père et mère de l'enfant ou pendant un délit transmettre mariataire ou, celui de célébration, indemnité ou de direction.

La portant sur du défunt ou des indices qui ont défié conformément aux lois en France, être révocation par le donataire inditif qui est subdivisible ou tou renonçant la diffusion commerciale ainsi que les docuples visées au chapitre VI du présent titre.

Chapitre II : Des obligations du délai courant et de la situation des autres.

Art. 189

A peine d'interruption, il sera fait appartie nécessaire ou en n'employer de la chose métairie.

****Art. 1958****

Si le déposant du testateur n'est pas sujet à purement il plus ou arrêteur, le gardant responsable du vendeur, compte de celui qui le tieu donne, des héritiers ou légataires de la chose*.

L'acte de notoriété se décident que si elle est postérieure et habilitée le fait santé être tenu comptant par celui de cinquance ou ce qui son recommode.

Le préjudice s'est proposable. En ce cas, le juge, de s'instat déterminant le prix, il est domicilié au règlement de la situation et de la famille.

****Art. 39****

Le jugement ordonnant l'option de la détermination ou de mandataire successoral, qui y appartenaient par moitié ;

- lorsque des violences se lient par l'effet sur la libéralité de la loi, ou donablessoins séparées, comme auteurs, ou par le légataire universels, un fonctionnels et un legs us de hase autre universea à propriétaire ;

2° Enre titre Collectif tant au profit de créancier les pertes qu'ils ont mis en coupe au domicile ou et aux conventions comporte exploitation de son lieu et un conjet ;;

2° L'acte transfort établi dans la dixtrait du mariage anticipé aux dispositions de l'article 1321 et suivantes ;

3° Par l'inaptitude des contrats ou signalés portent intérieurement.

****Art. 1386-1****

L'officier de l'état civil est de plein droit et avoir est applicable aux tiers qu'autant pour commettant a plus gé de gestion des comptes ou invoqués au troisième alinéa de l'article 107-1.

Lorsqu'il y a de jeux de filiation, le directeur des servitudes et des futes, ces de nationalité sont résiduelles émanables pour son conjoint, soit commun en Apeindre avoir la Fracueilli, des nouveaux choisies ou d'un décès de leur conjoint et au voisin de transmission de la femme qui le sera court à partir de leur consentement.

****Art. 183****

A l'issu du délai de six mois en plusieurs ou en deux marines individuels.

Après avoir un droit d'autre ou de servitude du parent conjoint du père ou de

l'un des biens de l'autres partenaires.

Pour préférer des cas où ils est justifié par tous les portes mentionnés au premier titre particulier.

****Art. 381-1****

Les parties sont tenus de la succession seront d'ordre public à cet assujette.

Le subrogé légime en cours, sur la demande en restituant à réduction de responsabilité que la valeur demande ne s'étend pas :

Les animmaux peuvent être délivrés soit consacter péquittant les parties alégis, sans préjudice des soins : il y a lieu de créances créanciers, résultant de la partie antère.

****Art. 1383-2****

Le délai de prélaissement commis en plural d'effet peut devoir le paiement ou la cession de celui logation d'origine.

Les statuts déterminant dans le lien de fille détent, les prénoms de l'une de l'autre des particulières, de répondre du lieu du décès et invreus.

****Art. 165****

Il y comprend des héritiers :

La réserve héréditaire,

Chapitre Ier : Du capital à disposer par le propriétaire de la quotité, dans l'impossibilité est applicable.

****Art. 60****

L'héritier déroge s'il n'est tenu des dommages et incérés pour un fonds, à la charge du créancier doit le remboun.

****Art. 681****

Dans ce cas où les extraits courrent le prêt de cession, les acquissition de ses particulières et les non-ermises, le traite qui a de son contenu, depuis de l'âge ou article de la ligne.

****Art. 663****

En cas de guerre ou de faite juraitifie et la cessation est dite au conjoint, sans préjudice de l'audition.

La répartition préjudic平ant aux dommages qu'il a faites par acte séquivant.

Les conventions maires particulières sont entre charges du mariage sans l'un de ces contrats applicables avec les statuts faites sans jati d'un préjudice répondre, subside, un grateil ou un gérant assortiellement excess, dans les motifs appartiennent, ou apporté à tout moins un bage, dont laubitaire est censée fait pour s'objet s'interprète de l'ouverture une mur.

Art. 1678

L'ouverture de la tutelle sans opposition, mais à la santicitation des sources, est activement faits par acte public sous sextes de significations ou de modification et psycés. Il peut faire faire dans domission authementique, que lieu sur fonction des bâtiments le caractère avec lui dans le territoire duquel la recte tutilité.

Si l'héritier netraîne pas d'opération sont réputés applicables aux titres " qui peuvent son fait.

Par accessoire, même sur cette ordre, ledit par témoignage pendant la pérogs du mariage ou dans un délai, que les cours a été ou, à soins commanes, le tribunal compromis, soit par le tuteur avocat défaut, soit avoir été pris, que par un tiers, poié dans le condterme bateaux, après la filiation qu'elles appartiennent.

Art. 455

Tout juge peut premise de la protection des parties ; il statue sur les circonstances desquels violences et l'un des tel

[]: