Fundations of Generative AI

Text Classification

Data Trainers
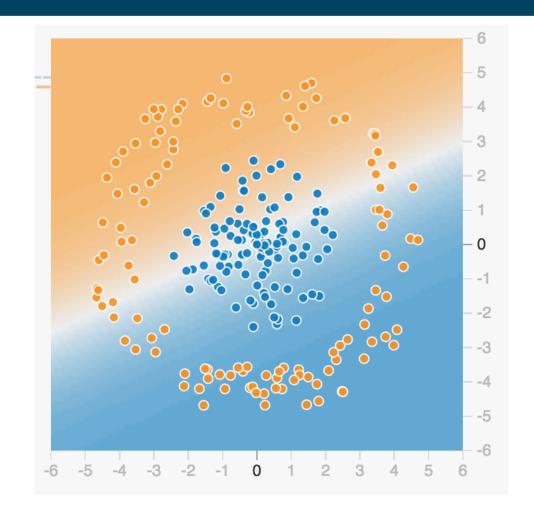
Foundations of
GenAI

Feed forward

Networks

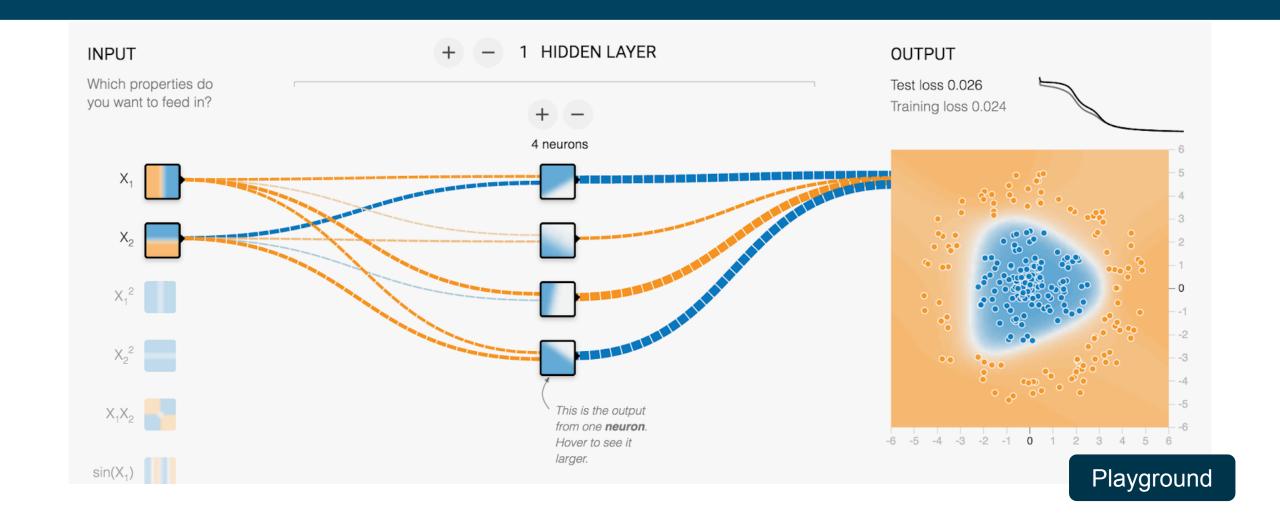▶ Can we separate this points to
Classify correctly the class
(seniment, category, etc… ?)

Playground
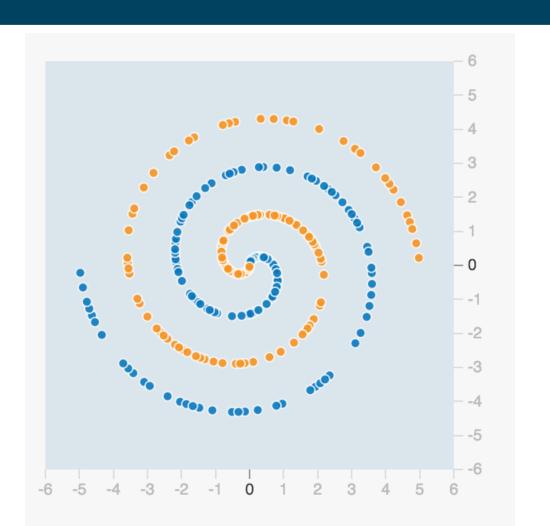
More neurons == nonlinearities

Playground

How can we do to split these?

Again, mini-lines help!

Data Trainers
Text Classification
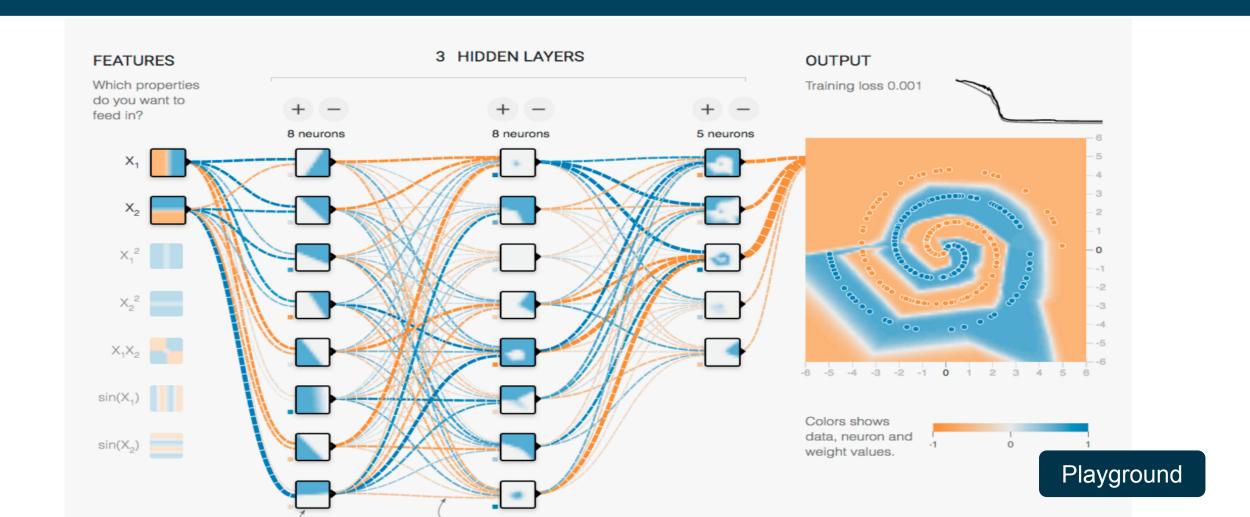
In code

Define a Sequential model

```
model = nn.Sequential(
    nn.Embedding(vocab_size, embedding_dim),
    nn.Linear(embedding_dim, 100),
    nn.ReLU(),
    nn.Linear(100, 50),
    nn.ReLU(),
    MeanLayer(),
    nn.Linear(50, 4),
    nn.Softmax()
)
```

In code

Add an embedding layer to the model

```python
model = nn.Sequential(
    nn.Embedding(vocab_size, embedding_dim),
    nn.Linear(embedding_dim, 100),
    nn.ReLU(),
    nn.Linear(100, 50),
    nn.ReLU(),
    MeanLayer(),
    nn.Linear(50, 4),
    nn.Softmax()
)
```

In code

Add a fully connected layer of 100 hidden units

```python
model = nn.Sequential(
    nn.Embedding(vocab_size, embedding_dim),
    nn.Linear(embedding_dim, 100),
    nn.ReLU(),
    nn.Linear(100, 50),
    nn.ReLU(),
    MeanLayer(),
    nn.Linear(50, 4),
    nn.Softmax()
)
```

Add the activation

```
model = nn.Sequential(
    nn.Embedding(vocab_size, embedding_dim),
    nn.Linear(embedding_dim, 100),
    nn.ReLU(),
    nn.Linear(100, 50),
    nn.ReLU(),
    MeanLayer(),
    nn.Linear(50, 4),
    nn.Softmax()
)
```

In code

Average across words

```
model = nn.Sequential(
    nn.Embedding(vocab_size, embedding_dim),
    nn.Linear(embedding_dim, 100),
    nn.ReLU(),
    nn.Linear(100, 50),
    nn.ReLU(),
    MeanLayer(),
    nn.Linear(50, 4),
    nn.Softmax()
)
```

In code

Final softmax
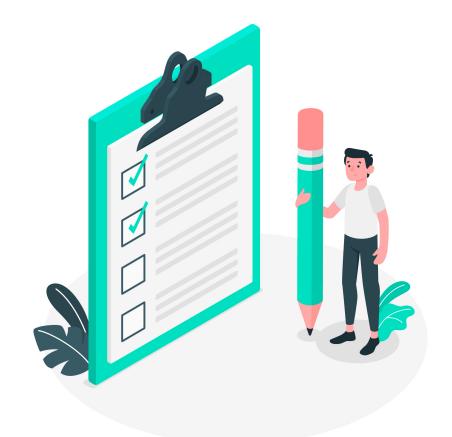
```
model = nn.Sequential(
    nn.Embedding(vocab_size, embedding_dim),
    nn.Linear(embedding_dim, 100),
    nn.ReLU(),
    nn.Linear(100, 50),
    nn.ReLU(),
    MeanLayer(),
    nn.Linear(50, 4),
    nn.Softmax()
)
```

LAB

News Classifier with NLP

▶ Use a feed forward neura

network to classify news

headlines!

- Adding more layers to neural networks results in a nonlinear decision boundary.

- This way we can make classification on text.

- FFN have ton of parameters, so they tend to not scale when the vocal is big, or the training set is complex.

- One thing you can do as homework is add fine-tuning of a pertained embedding layer to imporve results!