Pretzels & Tacos

Step 1: Creating a Source of Truth
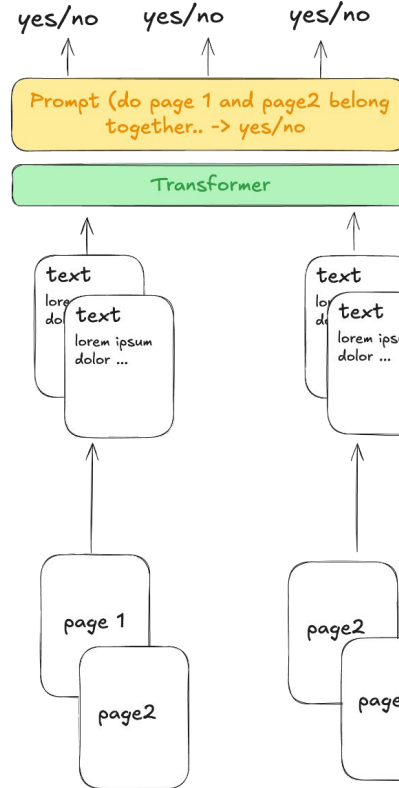
Idea 1: Text based Approaches
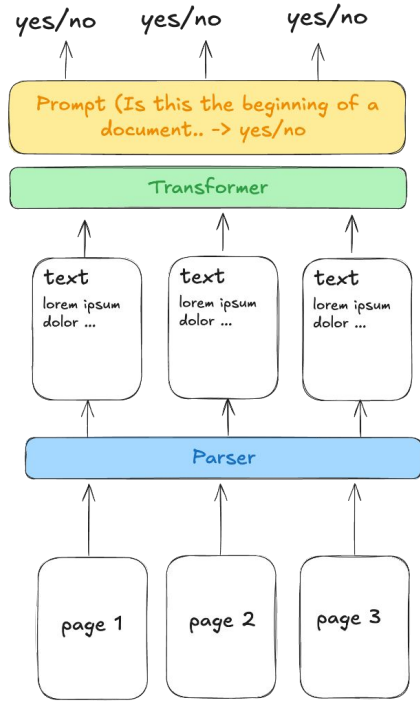
yes/no   yes/no   yes/no

Prompt (Is this the beginning of a document.. -> yes/no)

Transformer

text
lorem ipsum dolor ...

text
lorem ipsum dolor ...

text
lorem ipsum dolor ...

Parser

page 1

page 2

page 3

Idea:
Few-Shot give some examples with yes and no

Maybe only look at the first tokens

yes/no   yes/no   yes/no

Prompt (do page 1 and page2 belong together.. -> yes/no)

Transformer

text
lorem dol...

text
lorem ipsum dolor ...

text
lorem dol...

text
lorem ipsum dolor ...

page 1

page2

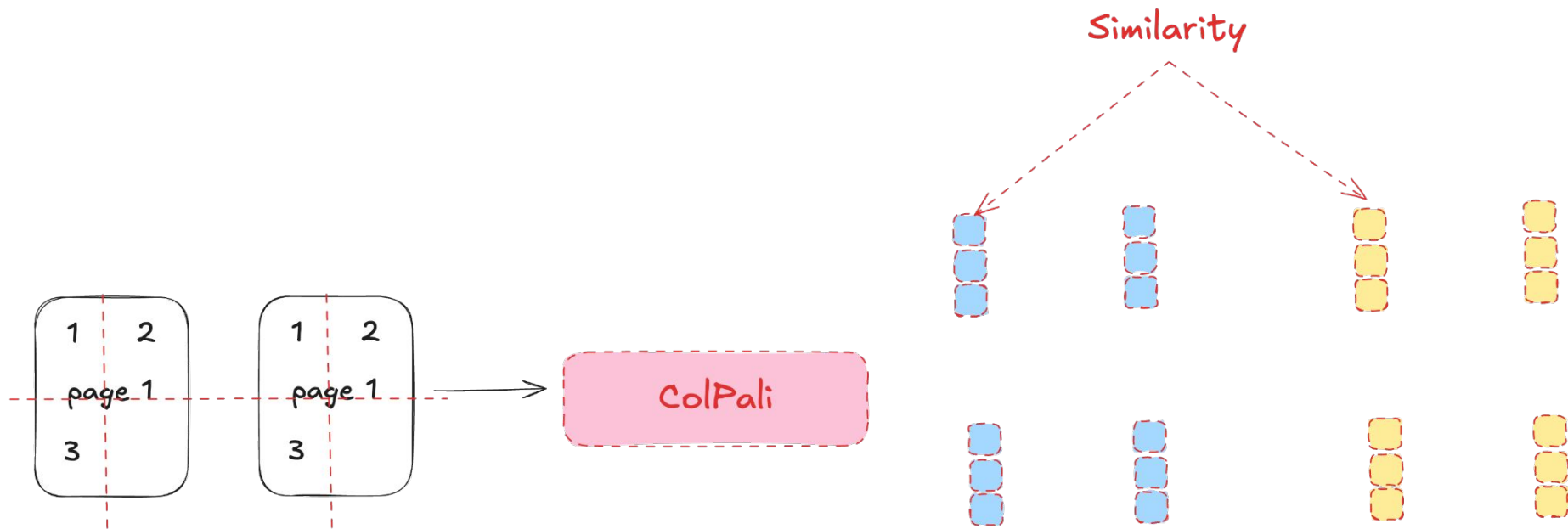page2

page 3

Idea:
Few-Shot give some examples with yes and no

Maybe only look at the first tokens

Idea 2: Vision Language Models

Idea 2: Vision Language Models