

## Reading from the text file

```
with open("D:/Users/USER/Documents/Build LLM from Scratch/the-verdict.txt", "r", encoding="utf-8") as f:
    raw_text = f.read()
print("Total number of character:", len(raw_text))

# next line is the code for printing first 100 characters
print(raw_text[:99])
```

Total number of character: 20479

I HAD always thought Jack Gisburn rather a cheap genius--though a good fellow enough--so it was no

Below code splits the text on spaces

```
import re
text = "Hello, world. This is test"

# r is added to consider the \ as literal characters
# \s captures the space whether it appears
# so basically we are asking python to keep a space while splitting
the text based on spaces.

result = re.split(r'(\s)', text)
print(result)

['Hello,', ' ', 'world.', ' ', 'This', ' ', 'is', ' ', 'test']
```

So far, we have splitted the text based on space..but the text also has some special character. The code below tries to split on those special characters

```
import re
text = "Hello, world. This is test"

# [,.] matches comma and a period
# | will look for either of any characters (comma or a period)

result = re.split(r'([,.]|\s)', text)
print(result)

['Hello', ',', ' ', 'world', '.', ' ', 'This', ' ', 'is', ' ', 'test']
```

The above result still has whitespaces which can be removed using the following code

```
result = [item for item in result if item.strip()]
print(result)

['Hello', ',', 'world', '.', 'This', 'is', 'test']
```

We need to make the code handle different other type of punctuations like question mark, quotation marks, double dashes.

```
text = "Hello, world. Is this -- a test?"
result = re.split(r'([.,:;?_!"()\`]|--|\s)', text)
result = [item.strip() for item in result if item.strip()]
print(result)

['Hello', ',', 'world', '.', 'Is', 'this', '--', 'a', 'test', '?']
```

## Applying the splitting done so far to 'The Verdict' text

```
preprocessed = re.split(r'([.,:;?_!"()\`]|--|\s)', raw_text)
preprocessed = [item.strip() for item in preprocessed if item.strip()]
print(preprocessed)

['I', 'HAD', 'always', 'thought', 'Jack', 'Gisburn', 'rather', 'a',
'cheap', 'genius', '--', 'though', 'a', 'good', 'fellow', 'enough',
'--', 'so', 'it', 'was', 'no', 'great', 'surprise', 'to', 'me', 'to',
'hear', 'that', ',', 'in', 'the', 'height', 'of', 'his', 'glory', ',',
'he', 'had', 'dropped', 'his', 'painting', ',', 'married', 'a',
'rich', 'widow', ',', 'and', 'established', 'himself', 'in', 'a',
'villa', 'on', 'the', 'Riviera', '.', '(', 'Though', 'I', 'rather',
'thought', 'it', 'would', 'have', 'been', 'Rome', 'or', 'Florence',
',', ')', '"', 'The', 'height', 'of', 'his', 'glory', '"', '--',
'that', 'was', 'what', 'the', 'women', 'called', 'it', '.', 'I',
'can', 'hear', 'Mrs', '.', 'Gideon', 'Thwing', '--', 'his', 'last',
'Chicago', 'sitter', '--', 'deploring', 'his', 'unaccountable',
'abdication', ',', '"', 'Of', 'course', 'it', '"', 's', 'going', 'to',
'send', 'the', 'value', 'of', 'my', 'picture', '"', 'way', 'up', ';',
'but', 'I', 'don', '"', 't', 'think', 'of', 'that', ',', 'Mr', '.',
'Rickham', '--', 'the', 'loss', 'to', 'Arret', 'is', 'all', 'I',
'think', 'of', '.', '"', 'The', 'word', ',', 'on', 'Mrs', '.',
'Thwing', '"', 's', 'lips', ',', 'multiplied', 'its', '-', 'rs', '-',
'as', 'though', 'they', 'were', 'reflected', 'in', 'an', 'endless',
'vista', 'of', 'mirrors', '.', 'And', 'it', 'was', 'not', 'only',
'the', 'Mrs', '.', 'Thwings', 'who', 'mourned', '.', 'Had', 'not',
'the', 'exquisite', 'Hermia', 'Croft', ',', 'at', 'the', 'last',
'Grafton', 'Gallery', 'show', ',', 'stopped', 'me', 'before',
'Gisburn', '"', 's', '"', 'Moon-dancers', '"', 'to', 'say', ',',
'with', 'tears', 'in', 'her', 'eyes', ':', '"', 'We', 'shall', 'not',
'look', 'upon', 'its', 'like', 'again', '"', '?', 'Well', '!', '--',
'even', 'through', 'the', 'prism', 'of', 'Hermia', '"', 's', 'tears',
'I', 'felt', 'able', 'to', 'face', 'the', 'fact', 'with',
'equanimity', '.', 'Poor', 'Jack', 'Gisburn', '!', 'The', 'women',
'had', 'made', 'him', '--', 'it', 'was', 'fitting', 'that', 'they',
'should', 'mourn', 'him', '.', 'Among', 'his', 'own', 'sex', 'fewer',
'regrets', 'were', 'heard', ',', 'and', 'in', 'his', 'own', 'trade',
'hardly', 'a', 'murmur', '.', 'Professional', 'jealousy', '?',
'Perhaps', '.', 'If', 'it', 'were', ',', 'the', 'honour', 'of', 'the',
```

'craft', 'was', 'vindicated', 'by', 'little', 'Claude', 'Nutley', ' ', 'who', ' ', 'in', 'all', 'good', 'faith', ' ', 'brought', 'out', 'in', 'the', 'Burlington', 'a', 'very', 'handsome', ' ', 'obituary', ' ', 'on', 'Jack', ' ', 'one', 'of', 'those', 'showy', 'articles', 'stocked', 'with', 'random', 'technicalities', 'that', 'I', 'have', 'heard', '(', 'I', 'won', ' ', 't', 'say', 'by', 'whom', ')', 'compared', 'to', 'Gisburn', ' ', 's', 'painting', ' ', 'And', 'so', ' ', 'his', 'resolve', 'being', 'apparently', 'irrevocable', ' ', 'the', 'discussion', 'gradually', 'died', 'out', ' ', 'and', ' ', 'as', 'Mrs', ' ', 'Thwing', 'had', 'predicted', ' ', 'the', 'price', 'of', ' ', 'Gisburns', ' ', 'went', 'up', ' ', 'It', 'was', 'not', 'till', 'three', 'years', 'later', 'that', ' ', 'in', 'the', 'course', 'of', 'a', 'few', 'weeks', ' ', 'idling', 'on', 'the', 'Riviera', ' ', 'it', 'suddenly', 'occurred', 'to', 'me', 'to', 'wonder', 'why', 'Gisburn', 'had', 'given', 'up', 'his', 'painting', ' ', 'On', 'reflection', ' ', 'it', 'really', 'was', 'a', 'tempting', 'problem', ' ', 'To', 'accuse', 'his', 'wife', 'would', 'have', 'been', 'too', 'easy', ' ', 'his', 'fair', 'sitters', 'had', 'been', 'denied', 'the', 'solace', 'of', 'saying', 'that', 'Mrs', ' ', 'Gisburn', 'had', ' ', 'dragged', 'him', 'down', ' ', 'For', 'Mrs', ' ', 'Gisburn', ' ', 'as', 'such', ' ', 'had', 'not', 'existed', 'till', 'nearly', 'a', 'year', 'after', 'Jack', ' ', 's', 'resolve', 'had', 'been', 'taken', ' ', 'It', 'might', 'be', 'that', 'he', 'had', 'married', 'her', ' ', 'since', 'he', 'liked', 'his', 'ease', ' ', 'because', 'he', 'didn', ' ', 't', 'want', 'to', 'go', 'on', 'painting', ' ', 'but', 'it', 'would', 'have', 'been', 'hard', 'to', 'prove', 'that', 'he', 'had', 'given', 'up', 'his', 'painting', 'because', 'he', 'had', 'married', 'her', ' ', 'Of', 'course', ' ', 'if', 'she', 'had', 'not', 'dragged', 'him', 'down', ' ', 'she', 'had', 'equally', ' ', 'as', 'Miss', 'Croft', 'contended', ' ', 'failed', 'to', ' ', 'lift', 'him', 'up', ' ', 'she', 'had', 'not', 'led', 'him', 'back', 'to', 'the', 'easel', ' ', 'To', 'put', 'the', 'brush', 'into', 'his', 'hand', 'again', ' ', 'what', 'a', 'vocation', 'for', 'a', 'wife', '!', 'But', 'Mrs', ' ', 'Gisburn', 'appeared', 'to', 'have', 'disdained', 'it', ' ', 'and', 'I', 'felt', 'it', 'might', 'be', 'interesting', 'to', 'find', 'out', 'why', ' ', 'The', 'desultory', 'life', 'of', 'the', 'Riviera', 'lends', 'itself', 'to', 'such', 'purely', 'academic', 'speculations', ' ', 'and', 'having', ' ', 'on', 'my', 'way', 'to', 'Monte', 'Carlo', ' ', 'caught', 'a', 'glimpse', 'of', 'Jack', ' ', 's', 'balustraded', 'terraces', 'between', 'the', 'pines', ' ', 'I', 'had', 'myself', 'borne', 'thither', 'the', 'next', 'day', ' ', 'I', 'found', 'the', 'couple', 'at', 'tea', 'beneath', 'their', 'palm-trees', ' ', 'and', 'Mrs', ' ', 'Gisburn', ' ', 's', 'welcome', 'was', 'so', 'genial', 'that', ' ', 'in', 'the', 'ensuing', 'weeks', ' ', 'I', 'claimed', 'it', 'frequently', ' ', 'It', 'was', 'not', 'that', 'my', 'hostess', 'was', ' ', 'interesting', ' ', 'on', 'that', 'point', 'I', 'could', 'have', 'given', 'Miss', 'Croft', 'the', 'fullest', 'reassurance', ' ', 'It', 'was', 'just', 'because', 'she', 'was', ' ', 'not', ' ', 'interesting', ' ', 'if', 'I', 'may',

'be', 'pardoned', 'the', 'bull', '--', 'that', 'I', 'found', 'her',  
'so', '.', 'For', 'Jack', ',', 'all', 'his', 'life', ',', 'had',  
'been', 'surrounded', 'by', 'interesting', 'women', ':', 'they',  
'had', 'fostered', 'his', 'art', ',', 'it', 'had', 'been', 'reared',  
'in', 'the', 'hot-house', 'of', 'their', 'adulation', '.', 'And',  
'it', 'was', 'therefore', 'instructive', 'to', 'note', 'what',  
'effect', 'the', '"', 'deadening', 'atmosphere', 'of', 'mediocrity',  
'"', '(', 'I', 'quote', 'Miss', 'Croft', ')', 'was', 'having', 'on',  
'him', '.', 'I', 'have', 'mentioned', 'that', 'Mrs', '.', 'Gisburn',  
'was', 'rich', ';', 'and', 'it', 'was', 'immediately', 'perceptible',  
'that', 'her', 'husband', 'was', 'extracting', 'from', 'this',  
'circumstance', 'a', 'delicate', 'but', 'substantial', 'satisfaction',  
'.'. 'It', 'is', ',', 'as', 'a', 'rule', ',', 'the', 'people', 'who',  
'scorn', 'money', 'who', 'get', 'most', 'out', 'of', 'it', ';', 'and',  
'Jack', '"', 's', 'elegant', 'disdain', 'of', 'his', 'wife', '"', 's',  
'big', 'balance', 'enabled', 'him', ',', 'with', 'an', 'appearance',  
'of', 'perfect', 'good-breeding', ',', 'to', 'transmute', 'it',  
'into', 'objects', 'of', 'art', 'and', 'luxury', '.', 'To', 'the',  
'latter', ',', 'I', 'must', 'add', ',', 'he', 'remained',  
'relatively', 'indifferent', ';', 'but', 'he', 'was', 'buying',  
'Renaissance', 'bronzes', 'and', 'eighteenth-century', 'pictures',  
'with', 'a', 'discrimination', 'that', 'bespoke', 'the', 'amplest',  
'resources', '.', '"', 'Money', '"', 's', 'only', 'excuse', 'is',  
'to', 'put', 'beauty', 'into', 'circulation', ',', '"', 'was', 'one',  
'of', 'the', 'axioms', 'he', 'laid', 'down', 'across', 'the',  
'Sevres', 'and', 'silver', 'of', 'an', 'exquisitely', 'appointed',  
'luncheon-table', ',', 'when', ',', 'on', 'a', 'later', 'day', ',',  
'I', 'had', 'again', 'run', 'over', 'from', 'Monte', 'Carlo', ';',  
'and', 'Mrs', '.', 'Gisburn', ',', 'beaming', 'on', 'him', ',',  
'added', 'for', 'my', 'enlightenment', ':', '"', 'Jack', 'is', 'so',  
'morbidly', 'sensitive', 'to', 'every', 'form', 'of', 'beauty', '.',  
'"', 'Poor', 'Jack', '!', 'It', 'had', 'always', 'been', 'his',  
'fate', 'to', 'have', 'women', 'say', 'such', 'things', 'of', 'him',  
'.:', 'the', 'fact', 'should', 'be', 'set', 'down', 'in',  
'extenuation', '.', 'What', 'struck', 'me', 'now', 'was', 'that', ',',  
'for', 'the', 'first', 'time', ',', 'he', 'resented', 'the', 'tone',  
'.'. 'I', 'had', 'seen', 'him', ',', 'so', 'often', ',', 'basking',  
'under', 'similar', 'tributes', '--', 'was', 'it', 'the', 'conjugal',  
'note', 'that', 'robbed', 'them', 'of', 'their', 'savour', '?', 'No',  
'--', 'for', ',', 'oddly', 'enough', ',', 'it', 'became', 'apparent',  
'that', 'he', 'was', 'fond', 'of', 'Mrs', '.', 'Gisburn', '--',  
'fond', 'enough', 'not', 'to', 'see', 'her', 'absurdity', '.', 'It',  
'was', 'his', 'own', 'absurdity', 'he', 'seemed', 'to', 'be',  
'wincing', 'under', '--', 'his', 'own', 'attitude', 'as', 'an',  
'object', 'for', 'garlands', 'and', 'incense', '.', '"', 'My', 'dear',  
'.', 'since', 'I', '"', 've', 'chucked', 'painting', 'people', 'don',  
'"', 't', 'say', 'that', 'stuff', 'about', 'me', '--', 'they', 'say',  
'it', 'about', 'Victor', 'Grindle', ',', '"', 'was', 'his', 'only',  
'protest', ',', 'as', 'he', 'rose', 'from', 'the', 'table', 'and',

'strolled', 'out', 'onto', 'the', 'sunlit', 'terrace', '.', 'I',  
'glanced', 'after', 'him', ',', 'struck', 'by', 'his', 'last', 'word',  
'.', 'Victor', 'Grindle', 'was', ',', 'in', 'fact', ',', 'becoming',  
'the', 'man', 'of', 'the', 'moment', '--', 'as', 'Jack', 'himself',  
'.', 'one', 'might', 'put', 'it', ',', 'had', 'been', 'the', 'man',  
'of', 'the', 'hour', '.', 'The', 'younger', 'artist', 'was', 'said',  
'to', 'have', 'formed', 'himself', 'at', 'my', 'friend', '"', 's',  
'feet', ',', 'and', 'I', 'wondered', 'if', 'a', 'tinge', 'of',  
'jealousy', 'underlay', 'the', 'latter', '"', 's', 'mysterious',  
'abdication', '.', 'But', 'no', '--', 'for', 'it', 'was', 'not',  
'till', 'after', 'that', 'event', 'that', 'the', ' ', 'rose',  
'Dubarry', ' ', 'drawing-rooms', 'had', 'begun', 'to', 'display',  
'their', '"', 'Grindles', '.', '"', 'I', 'turned', 'to', 'Mrs', '.',  
'Gisburn', ',', 'who', 'had', 'lingered', 'to', 'give', 'a', 'lump',  
'of', 'sugar', 'to', 'her', 'spaniel', 'in', 'the', 'dining-room',  
'.', '"', 'Why', ' ', 'has', ' ', 'he', 'chucked', 'painting', '?',  
'"', 'I', 'asked', 'abruptly', '.', 'She', 'raised', 'her',  
'eyebrows', 'with', 'a', 'hint', 'of', 'good-humoured', 'surprise',  
'.', '"', 'Oh', ' ', 'he', 'doesn', '"', 't', ' ', 'have', ' ', 'to',  
'now', ',', 'you', 'know', ';;', 'and', 'I', 'want', 'him', 'to',  
'enjoy', 'himself', ',', '"', 'she', 'said', 'quite', 'simply', '.',  
'I', 'looked', 'about', 'the', 'spacious', 'white-panelled', 'room',  
' ', 'with', 'its', ' ', 'famille-verte', ' ', 'vases', 'repeating',  
'the', 'tones', 'of', 'the', 'pale', 'damask', 'curtains', ',', 'and',  
'its', 'eighteenth-century', 'pastels', 'in', 'delicate', 'faded',  
'frames', '.', '"', 'Has', 'he', 'chucked', 'his', 'pictures', 'too',  
'?', 'I', 'haven', '"', 't', 'seen', 'a', 'single', 'one', 'in',  
'the', 'house', '.', '"', 'A', 'slight', 'shade', 'of', 'constraint',  
'crossed', 'Mrs', '.', 'Gisburn', '"', 's', 'open', 'countenance',  
'.', '"', 'It', '"', 's', 'his', 'ridiculous', 'modesty', ',', 'you',  
'know', '.', 'He', 'says', 'they', '"', 're', 'not', 'fit', 'to',  
'have', 'about', ';;', 'he', '"', 's', 'sent', 'them', 'all', 'away',  
'except', 'one', '--', 'my', 'portrait', '--', 'and', 'that', 'I',  
'have', 'to', 'keep', 'upstairs', '.', '"', 'His', 'ridiculous',  
'modesty', '--', 'Jack', '"', 's', 'modesty', 'about', 'his',  
'pictures', '?', 'My', 'curiosity', 'was', 'growing', 'like', 'the',  
'bean-stalk', '.', 'I', 'said', 'persuasively', 'to', 'my', 'hostess',  
'.', '"', 'I', 'must', 'really', 'see', 'your', 'portrait', ',',  
'you', 'know', '.', '"', 'She', 'glanced', 'out', 'almost',  
'timorously', 'at', 'the', 'terrace', 'where', 'her', 'husband', ' ',  
'lounging', 'in', 'a', 'hooded', 'chair', ',', 'had', 'lit', 'a',  
'cigar', 'and', 'drawn', 'the', 'Russian', 'deerhound', '"', 's',  
'head', 'between', 'his', 'knees', '.', '"', 'Well', ',', 'come',  
'while', 'he', '"', 's', 'not', 'looking', ',', '"', 'she', 'said',  
' ', 'with', 'a', 'laugh', 'that', 'tried', 'to', 'hide', 'her',  
'nervousness', ';;', 'and', 'I', 'followed', 'her', 'between', 'the',  
'marble', 'Emperors', 'of', 'the', 'hall', ',', 'and', 'up', 'the',  
'wide', 'stairs', 'with', 'terra-cotta', 'nymphs', 'poised', 'among',  
'flowers', 'at', 'each', 'landing', '.', 'In', 'the', 'dimpest',

'corner', 'of', 'her', 'boudoir', ',', 'amid', 'a', 'profusion', 'of',  
'delicate', 'and', 'distinguished', 'objects', ',', 'hung', 'one',  
'of', 'the', 'familiar', 'oval', 'canvases', ',', 'in', 'the',  
'inevitable', 'garlanded', 'frame', '.', 'The', 'mere', 'outline',  
'of', 'the', 'frame', 'called', 'up', 'all', 'Gisburn', '"', 's',  
'past', '!', 'Mrs', '.', 'Gisburn', 'drew', 'back', 'the', 'window-  
curtains', ',', 'moved', 'aside', 'a', 'jardiniere', 'full',  
'of', 'pink', 'azaleas', ',', 'pushed', 'an', 'arm-chair', 'away',  
', 'and', 'said', ':', '"', 'If', 'you', 'stand', 'here', 'you',  
'can', 'just', 'manage', 'to', 'see', 'it', '.', 'I', 'had', 'it',  
'over', 'the', 'mantel-piece', ',', 'but', 'he', 'wouldn', '"', 't',  
'let', 'it', 'stay', '.', '"', 'Yes', '--', 'I', 'could', 'just',  
'manage', 'to', 'see', 'it', '--', 'the', 'first', 'portrait', 'of',  
'Jack', '"', 's', 'I', 'had', 'ever', 'had', 'to', 'strain', 'my',  
'eyes', 'over', '!', 'Usually', 'they', 'had', 'the', 'place', 'of',  
'honour', '--', 'say', 'the', 'central', 'panel', 'in', 'a', 'pale',  
'yellow', 'or', 'a', 'rose', 'Dubarry', 'drawing-room',  
'or', 'a', 'monumental', 'easel', 'placed', 'so', 'that', 'it',  
'took', 'the', 'light', 'through', 'curtains', 'of', 'old',  
'Venetian', 'point', '.', 'The', 'more', 'modest', 'place', 'became',  
'the', 'picture', 'better', ';', 'yet', ',', 'as', 'my', 'eyes',  
'grew', 'accustomed', 'to', 'the', 'half-light', ',', 'all', 'the',  
'characteristic', 'qualities', 'came', 'out', '--', 'all', 'the',  
'hesitations', 'disguised', 'as', 'audacities', ',', 'the', 'tricks',  
'of', 'prestidigitation', 'by', 'which', ',', 'with', 'such',  
'consummate', 'skill', ',', 'he', 'managed', 'to', 'divert',  
'attention', 'from', 'the', 'real', 'business', 'of', 'the',  
'picture', 'to', 'some', 'pretty', 'irrelevance', 'of', 'detail', '.',  
'Mrs', '.', 'Gisburn', ',', 'presenting', 'a', 'neutral', 'surface',  
'to', 'work', 'on', '--', 'forming', ',', 'as', 'it', 'were', ',',  
'so', 'inevitably', 'the', 'background', 'of', 'her', 'own',  
'picture', '--', 'had', 'lent', 'herself', 'in', 'an', 'unusual',  
'degree', 'to', 'the', 'display', 'of', 'this', 'false', 'virtuosity',  
', 'The', 'picture', 'was', 'one', 'of', 'Jack', '"', 's', '"',  
'strongest', ',', '"', 'as', 'his', 'admirers', 'would', 'have',  
'put', 'it', '--', 'it', 'represented', ',', 'on', 'his', 'part', ',',  
'a', 'swelling', 'of', 'muscles', ',', 'a', 'congesting', 'of',  
'veins', ',', 'a', 'balancing', ',', 'straddling', 'and', 'straining',  
', 'that', 'reminded', 'one', 'of', 'the', 'circus-clown', '"', 's',  
'ironic', 'efforts', 'to', 'lift', 'a', 'feather', '.', 'It', 'met',  
', 'in', 'short', ',', 'at', 'every', 'point', 'the', 'demand',  
'of', 'lovely', 'woman', 'to', 'be', 'painted', '"', 'strongly', '"',  
'because', 'she', 'was', 'tired', 'of', 'being', 'painted', '"',  
'sweetly', '"', '--', 'and', 'yet', 'not', 'to', 'lose', 'an', 'atom',  
'of', 'the', 'sweetness', '.', '"', 'It', '"', 's', 'the', 'last',  
'he', 'painted', ',', 'you', 'know', ',', '"', 'Mrs', '.', 'Gisburn',  
'said', 'with', 'pardonable', 'pride', '.', '"', 'The', 'last', 'but',  
'one', ',', '"', 'she', 'corrected', 'herself', '--', '"', 'but',  
'the', 'other', 'doesn', '"', 't', 'count', ',', 'because', 'he',

'destroyed', 'it', '.', '"', '"', 'Destroyed', 'it', '?', '"', 'I',  
'was', 'about', 'to', 'follow', 'up', 'this', 'clue', 'when', 'I',  
'heard', 'a', 'footstep', 'and', 'saw', 'Jack', 'himself', 'on',  
'the', 'threshold', '.', 'As', 'he', 'stood', 'there', ',', 'his',  
'hands', 'in', 'the', 'pockets', 'of', 'his', 'velveteen', 'coat',  
',', 'the', 'thin', 'brown', 'waves', 'of', 'hair', 'pushed', 'back',  
'from', 'his', 'white', 'forehead', ',', 'his', 'lean', 'sunburnt',  
'cheeks', 'furrowed', 'by', 'a', 'smile', 'that', 'lifted', 'the',  
'tips', 'of', 'a', 'self-confident', 'moustache', ',', 'I', 'felt',  
'to', 'what', 'a', 'degree', 'he', 'had', 'the', 'same', 'quality',  
'as', 'his', 'pictures', '--', 'the', 'quality', 'of', 'looking',  
'cleverer', 'than', 'he', 'was', '.', 'His', 'wife', 'glanced', 'at',  
'him', 'deprecatingly', ',', 'but', 'his', 'eyes', 'travelled',  
'past', 'her', 'to', 'the', 'portrait', '.', '"', 'Mr', '.',  
'Rickham', 'wanted', 'to', 'see', 'it', ',', '"', 'she', 'began', ',',  
'as', 'if', 'excusing', 'herself', '.', 'He', 'shrugged', 'his',  
'shoulders', ',', 'still', 'smiling', '.', '"', 'Oh', ',', 'Rickham',  
'found', 'me', 'out', 'long', 'ago', ',', '"', 'he', 'said',  
'lightly', ';', 'then', ',', 'passing', 'his', 'arm', 'through',  
'mine', ':', '"', 'Come', 'and', 'see', 'the', 'rest', 'of', 'the',  
'house', '.', '"', 'He', 'showed', 'it', 'to', 'me', 'with', 'a',  
'kind', 'of', 'naive', 'suburban', 'pride', ':', 'the', 'bath-rooms',  
',', 'the', 'speaking-tubes', ',', 'the', 'dress-closets', ',', 'the',  
'trouser-presses', '--', 'all', 'the', 'complex', 'simplifications',  
'of', 'the', 'millionaire', '"', 's', 'domestic', 'economy', '.',  
'And', 'whenever', 'my', 'wonder', 'paid', 'the', 'expected',  
'tribute', 'he', 'said', ',', 'throwing', 'out', 'his', 'chest', 'a',  
'little', ':', '"', 'Yes', ',', 'I', 'really', 'don', '"', 't', 'see',  
'how', 'people', 'manage', 'to', 'live', 'without', 'that', '.', '"',  
'Well', '--', 'it', 'was', 'just', 'the', 'end', 'one', 'might',  
'have', 'foreseen', 'for', 'him', '.', 'Only', 'he', 'was', ',',  
'through', 'it', 'all', 'and', 'in', 'spite', 'of', 'it', 'all', '--',  
'as', 'he', 'had', 'been', 'through', ',', 'and', 'in', 'spite', 'of',  
',', 'his', 'pictures', '--', 'so', 'handsome', ',', 'so', 'charming',  
',', 'so', 'disarming', ',', 'that', 'one', 'longed', 'to', 'cry',  
'out', ':', '"', 'Be', 'dissatisfied', 'with', 'your', 'leisure', '!',  
'"', 'as', 'once', 'one', 'had', 'longed', 'to', 'say', ':', '"',  
'Be', 'dissatisfied', 'with', 'your', 'work', '!', '"', 'But', ',',  
'with', 'the', 'cry', 'on', 'my', 'lips', ',', 'my', 'diagnosis',  
'suffered', 'an', 'unexpected', 'check', '.', '"', 'This', 'is', 'my',  
'own', 'lair', ',', '"', 'he', 'said', ',', 'leading', 'me', 'into',  
'a', 'dark', 'plain', 'room', 'at', 'the', 'end', 'of', 'the',  
'florid', 'vista', '.', 'It', 'was', 'square', 'and', 'brown', 'and',  
'leathery', ':', 'no', '"', 'effects', '"', ';', 'no', 'bric-a-brac',  
',', 'none', 'of', 'the', 'air', 'of', 'posing', 'for',  
'reproduction', 'in', 'a', 'picture', 'weekly', '--', 'above', 'all',  
',', 'no', 'least', 'sign', 'of', 'ever', 'having', 'been', 'used',  
'as', 'a', 'studio', '.', 'The', 'fact', 'brought', 'home', 'to',  
'me', 'the', 'absolute', 'finality', 'of', 'Jack', '"', 's', 'break',

'with', 'his', 'old', 'life', '.', '"', 'Don', '"', 't', 'you',  
'ever', 'dabble', 'with', 'paint', 'any', 'more', '?', '"', 'I',  
'asked', ',', 'still', 'looking', 'about', 'for', 'a', 'trace', 'of',  
'such', 'activity', '.', '"', 'Never', ',', '"', 'he', 'said',  
'briefly', '.', '"', 'Or', 'water-colour', '--', 'or', 'etching', '?',  
'"', 'His', 'confident', 'eyes', 'grew', 'dim', ',', 'and', 'his',  
'cheeks', 'paled', 'a', 'little', 'under', 'their', 'handsome',  
'sunburn', '.', '"', 'Never', 'think', 'of', 'it', ',', 'my', 'dear',  
'fellow', '--', 'any', 'more', 'than', 'if', 'I', '"', 'd', 'never',  
'touched', 'a', 'brush', '.', '"', 'And', 'his', 'tone', 'told', 'me',  
'in', 'a', 'flash', 'that', 'he', 'never', 'thought', 'of',  
'anything', 'else', '.', 'I', 'moved', 'away', ',', 'instinctively',  
'embarrassed', 'by', 'my', 'unexpected', 'discovery', ';', 'and',  
'as', 'I', 'turned', ',', 'my', 'eye', 'fell', 'on', 'a', 'small',  
'picture', 'above', 'the', 'mantel-piece', '--', 'the', 'only',  
'object', 'breaking', 'the', 'plain', 'oak', 'panelling', 'of', 'the',  
'room', '.', '"', 'Oh', ',', 'by', 'Jove', '!', '"', 'I', 'said', '.',  
'It', 'was', 'a', 'sketch', 'of', 'a', 'donkey', '--', 'an', 'old',  
'tired', 'donkey', ',', 'standing', 'in', 'the', 'rain', 'under', 'a',  
'wall', '.', '"', 'By', 'Jove', '--', 'a', 'Stroud', '!', '"', 'I',  
'cried', '.', 'He', 'was', 'silent', ';', 'but', 'I', 'felt', 'him',  
'close', 'behind', 'me', ',', 'breathing', 'a', 'little', 'quickly',  
'"', 'What', 'a', 'wonder', '!', 'Made', 'with', 'a', 'dozen',  
'lines', '--', 'but', 'on', 'everlasting', 'foundations', '.', 'You',  
'lucky', 'chap', ',', 'where', 'did', 'you', 'get', 'it', '?', '"',  
'He', 'answered', 'slowly', ':', '"', 'Mrs', '.', 'Stroud', 'gave',  
'it', 'to', 'me', '.', '"', 'Ah', '--', 'I', 'didn', '"', 't',  
'know', 'you', 'even', 'knew', 'the', 'Strouds', '.', 'He', 'was',  
'such', 'an', 'inflexible', 'hermit', '.', '"', 'I', 'didn', '"',  
't', '--', 'till', 'after', ',', 'She', 'sent', 'for',  
'me', 'to', 'paint', 'him', 'when', 'he', 'was', 'dead', '.', '"',  
'"', 'When', 'he', 'was', 'dead', '?', 'You', '?', '"', 'I', 'must',  
'have', 'let', 'a', 'little', 'too', 'much', 'amazement', 'escape',  
'through', 'my', 'surprise', ',', 'for', 'he', 'answered', 'with',  
'a', 'deprecating', 'laugh', ':', '"', 'Yes', '--', 'she', '"', 's',  
'an', 'awful', 'simpleton', ',', 'you', 'know', ',', 'Mrs',  
'Stroud', '.', 'Her', 'only', 'idea', 'was', 'to', 'have', 'him',  
'done', 'by', 'a', 'fashionable', 'painter', '--', 'ah', ',', 'poor',  
'Stroud', '!', 'She', 'thought', 'it', 'the', 'surest', 'way', 'of',  
'proclaiming', 'his', 'greatness', '--', 'of', 'forcing', 'it', 'on',  
'a', 'purblind', 'public', '.', 'And', 'at', 'the', 'moment', 'I',  
'was', '—', 'the', '—', 'fashionable', 'painter', ',', '"', 'Ah',  
'—', 'poor', 'Stroud', '--', 'as', 'you', 'say', ',', 'Was',  
'that', '—', 'his', 'history', '?', '"', 'That', 'was', 'his',  
'history', '.', 'She', 'believed', 'in', 'him', ',', 'gloried', 'in',  
'him', '--', 'or', 'thought', 'she', 'did', '.', 'But', 'she',  
'couldn', '"', 't', 'bear', 'not', 'to', 'have', 'all', 'the',  
'drawing-rooms', 'with', 'her', '.', 'She', 'couldn', '"', 't',  
'bear', 'the', 'fact', 'that', '—', 'on', 'varnishing', 'days', '—',



'one', 'could', 'always', 'get', 'near', 'enough', 'to', 'see', 'his',  
'pictures', '...', 'Poor', 'woman', '!', 'She', '...', 's', 'just', 'a',  
'fragment', 'groping', 'for', 'other', 'fragments', '...', 'Stroud',  
'is', 'the', 'only', 'whole', 'I', 'ever', 'knew', '...', '...', '...',  
'You', 'ever', 'knew', '?', 'But', 'you', 'just', 'said', '...', '...',  
'Gisburn', 'had', 'a', 'curious', 'smile', 'in', 'his', 'eyes', '...',  
'...', 'Oh', '...', 'I', 'knew', 'him', '...', 'and', 'he', 'knew', 'me',  
'...', 'only', 'it', 'happened', 'after', 'he', 'was', 'dead', '...',  
'...', 'I', 'dropped', 'my', 'voice', 'instinctively', '...', '...', 'When',  
'she', 'sent', 'for', 'you', '?', '...', '...', 'Yes', '...', 'quite',  
'insensible', 'to', 'the', 'irony', '...', 'She', 'wanted', 'him',  
'vindicated', '...', 'and', 'by', 'me', '!', '...', 'He', 'laughed',  
'again', '...', 'and', 'threw', 'back', 'his', 'head', 'to', 'look',  
'up', 'at', 'the', 'sketch', 'of', 'the', 'donkey', '...', '...', 'There',  
'were', 'days', 'when', 'I', 'couldn', '...', 't', 'look', 'at', 'that',  
'thing', '...', 'couldn', '...', 't', 'face', 'it', '...', 'But', 'I',  
'forced', 'myself', 'to', 'put', 'it', 'here', '...', 'and', 'now',  
'it', '...', 's', 'cured', 'me', '...', 'cured', 'me', '...', 'That', '...',  
's', 'the', 'reason', 'why', 'I', 'don', '...', 't', 'dabble', 'any',  
'more', '...', 'my', 'dear', 'Rickham', '...', 'or', 'rather', 'Stroud',  
'himself', 'is', 'the', 'reason', '...', '...', 'For', 'the', 'first',  
'time', 'my', 'idle', 'curiosity', 'about', 'my', 'companion',  
'turned', 'into', 'a', 'serious', 'desire', 'to', 'understand', 'him',  
'better', '...', '...', 'I', 'wish', 'you', '...', 'd', 'tell', 'me', 'how',  
'it', 'happened', '...', '...', 'I', 'said', '...', 'He', 'stood',  
'looking', 'up', 'at', 'the', 'sketch', '...', 'and', 'twirling',  
'between', 'his', 'fingers', 'a', 'cigarette', 'he', 'had',  
'forgotten', 'to', 'light', '...', 'Suddenly', 'he', 'turned', 'toward',  
'me', '...', '...', 'I', '...', 'd', 'rather', 'like', 'to', 'tell', 'you',  
'...', 'because', 'I', '...', 've', 'always', 'suspected', 'you', 'of',  
'loathing', 'my', 'work', '...', '...', 'I', 'made', 'a', 'deprecating',  
'gesture', '...', 'which', 'he', 'negatived', 'with', 'a', 'good-  
humoured', 'shrug', '...', '...', 'Oh', '...', 'I', 'didn', '...', 't',  
'care', 'a', 'straw', 'when', 'I', 'believed', 'in', 'myself', '...',  
'and', 'now', 'it', '...', 's', 'an', 'added', 'tie', 'between', 'us',  
'!', '...', 'He', 'laughed', 'slightly', '...', 'without', 'bitterness',  
'...', 'and', 'pushed', 'one', 'of', 'the', 'deep', 'arm-chairs',  
'forward', '...', '...', 'There', '...', 'make', 'yourself', 'comfortable',  
'...', 'and', 'here', 'are', 'the', 'cigars', 'you', 'like', '...', '...',  
'He', 'placed', 'them', 'at', 'my', 'elbow', 'and', 'continued', 'to',  
'wander', 'up', 'and', 'down', 'the', 'room', '...', 'stopping', 'now',  
'and', 'then', 'beneath', 'the', 'picture', '...', '...', 'How', 'it',  
'happened', '?', 'I', 'can', 'tell', 'you', 'in', 'five', 'minutes',  
'...', 'and', 'it', 'didn', '...', 't', 'take', 'much', 'longer', 'to',  
'happen', '...', '...', '...', 'I', 'can', 'remember', 'now', 'how',  
'surprised', 'and', 'pleased', 'I', 'was', 'when', 'I', 'got', 'Mrs',  
'...', 'Stroud', '...', 's', 'note', '...', 'Of', 'course', '...', 'deep',  
'down', '...', 'I', 'had', 'always', '...', 'felt', '...', 'there', 'was',  
'no', 'one', 'like', 'him', '...', 'only', 'I', 'had', 'gone', 'with',

'the', 'stream', ',', 'echoed', 'the', 'usual', 'platitudes', 'about',  
'him', ',', 'till', 'I', 'half', 'got', 'to', 'think', 'he', 'was',  
'a', 'failure', ',', 'one', 'of', 'the', 'kind', 'that', 'are',  
'left', 'behind', ',', 'By', 'Jove', ',', 'and', 'he', '—', 'was',  
'—', 'left', 'behind', '—', 'because', 'he', 'had', 'come', 'to',  
'stay', '!', 'The', 'rest', 'of', 'us', 'had', 'to', 'let',  
'ourselves', 'be', 'swept', 'along', 'or', 'go', 'under', ',', 'but',  
'he', 'was', 'high', 'above', 'the', 'current', '—', 'on',  
'everlasting', 'foundations', ',', 'as', 'you', 'say', ',', '""',  
'Well', ',', 'I', 'went', 'off', 'to', 'the', 'house', 'in', 'my',  
'most', 'egregious', 'mood', '—', 'rather', 'moved', ',', 'Lord',  
'forgive', 'me', ',', 'at', 'the', 'pathos', 'of', 'poor', 'Stroud',  
'""', 's', 'career', 'of', 'failure', 'being', 'crowned', 'by', 'the',  
'glory', 'of', 'my', 'painting', 'him', '!', 'Of', 'course', 'I',  
'meant', 'to', 'do', 'the', 'picture', 'for', 'nothing', '—', 'I',  
'told', 'Mrs', ',', 'Stroud', 'so', 'when', 'she', 'began', 'to',  
'stammer', 'something', 'about', 'her', 'poverty', ',', 'I',  
'remember', 'getting', 'off', 'a', 'prodigious', 'phrase', 'about',  
'the', 'honour', 'being', '—', 'mine', '—', 'oh', ',', 'I',  
'was', 'princely', ',', 'my', '—', 'dear', 'Rickham', '!', 'I', 'was',  
'posing', 'to', 'myself', 'like', 'one', 'of', 'my', 'own', 'sitters',  
'—', '""', 'Then', 'I', 'was', 'taken', 'up', 'and', 'left', 'alone',  
'with', 'him', ',', 'I', 'had', 'sent', 'all', 'my', 'traps', 'in',  
'advance', ',', 'and', 'I', 'had', 'only', 'to', 'set', 'up', 'the',  
'easel', 'and', 'get', 'to', 'work', ',', 'He', 'had', 'been', 'dead',  
'only', 'twenty-four', 'hours', ',', 'and', 'he', 'died', 'suddenly',  
'—', 'of', 'heart', 'disease', ',', 'so', 'that', 'there', 'had',  
'been', 'no', 'preliminary', 'work', 'of', 'destruction', '—', 'his',  
'face', 'was', 'clear', 'and', 'untouched', ',', 'I', 'had', 'met',  
'him', 'once', 'or', 'twice', ',', 'years', 'before', ',', 'and',  
'thought', 'him', 'insignificant', 'and', 'dingy', ',', 'Now', 'I',  
'saw', 'that', 'he', 'was', 'superb', ',', '""', 'I', 'was', 'glad',  
'at', 'first', ',', 'with', 'a', 'merely', 'aesthetic',  
'satisfaction', ',', 'glad', 'to', 'have', 'my', 'hand', 'on', 'such',  
'a', '""', 'subject', ',', '""', 'Then', 'his', 'strange', 'life-  
likeness', 'began', 'to', 'affect', 'me', 'queerly', '—', 'as', 'I',  
'blocked', 'the', 'head', 'in', 'I', 'felt', 'as', 'if', 'he', 'were',  
'watching', 'me', 'do', 'it', ',', 'The', 'sensation', 'was',  
'followed', 'by', 'the', 'thought', ':', 'if', 'he', '—', 'were', '—',  
'watching', 'me', ',', 'what', 'would', 'he', 'say', 'to', 'my',  
'way', 'of', 'working', '?', 'My', 'strokes', 'began', 'to', 'go',  
'a', 'little', 'wild', '—', 'I', 'felt', 'nervous', 'and',  
'uncertain', ',', '""', 'Once', ',', 'when', 'I', 'looked', 'up', ',',  
'I', 'seemed', 'to', 'see', 'a', 'smile', 'behind', 'his', 'close',  
'grayish', 'beard', '—', 'as', 'if', 'he', 'had', 'the', 'secret',  
'—', 'and', 'were', 'amusing', 'himself', 'by', 'holding', 'it',  
'back', 'from', 'me', ',', 'That', 'exasperated', 'me', 'still',  
'more', ',', 'The', 'secret', '?', 'Why', ',', 'I', 'had', 'a',  
'secret', 'worth', 'twenty', 'of', 'his', '!', 'I', 'dashed', 'at',

'the', 'canvas', 'furiously', ',', 'and', 'tried', 'some', 'of', 'my',  
'bravura', 'tricks', '.', 'But', 'they', 'failed', 'me', ',', 'they',  
'crumbled', '.', 'I', 'saw', 'that', 'he', 'wasn', '"', 't',  
'watching', 'the', 'showy', 'bits', '--', 'I', 'couldn', '"', 't',  
'distract', 'his', 'attention', ';', 'he', 'just', 'kept', 'his',  
'eyes', 'on', 'the', 'hard', 'passages', 'between', '.', 'Those',  
'were', 'the', 'ones', 'I', 'had', 'always', 'shirked', ',', 'or',  
'covered', 'up', 'with', 'some', 'lying', 'paint', '.', 'And', 'how',  
'he', 'saw', 'through', 'my', 'lies', '!', '"', 'I', 'looked', 'up',  
'again', ',', 'and', 'caught', 'sight', 'of', 'that', 'sketch', 'of',  
'the', 'donkey', 'hanging', 'on', 'the', 'wall', 'near', 'his', 'bed',  
'.', 'His', 'wife', 'told', 'me', 'afterward', 'it', 'was', 'the',  
'last', 'thing', 'he', 'had', 'done', '--', 'just', 'a', 'note',  
'taken', 'with', 'a', 'shaking', 'hand', ',', 'when', 'he', 'was',  
'down', 'in', 'Devonshire', 'recovering', 'from', 'a', 'previous',  
'heart', 'attack', '.', 'Just', 'a', 'note', '!', 'But', 'it',  
'tells', 'his', 'whole', 'history', '.', 'There', 'are', 'years',  
'of', 'patient', 'scornful', 'persistence', 'in', 'every', 'line',  
'.', 'A', 'man', 'who', 'had', 'swum', 'with', 'the', 'current',  
'could', 'never', 'have', 'learned', 'that', 'mighty', 'up-stream',  
'stroke', ',', ',', ',', ',', ',', ',', 'I', 'turned', 'back', 'to', 'my',  
'work', ',', 'and', 'went', 'on', 'groping', 'and', 'muddling', ';',  
'then', 'I', 'looked', 'at', 'the', 'donkey', 'again', '.', 'I',  
'saw', 'that', ',', 'when', 'Stroud', 'laid', 'in', 'the', 'first',  
'stroke', ',', 'he', 'knew', 'just', 'what', 'the', 'end', 'would',  
'be', '.', 'He', 'had', 'possessed', 'his', 'subject', ',',  
'absorbed', 'it', ',', 'recreated', 'it', '.', 'When', 'had', 'I',  
'done', 'that', 'with', 'any', 'of', 'my', 'things', '?', 'They',  
'hadn', '"', 't', 'been', 'born', 'of', 'me', '--', 'I', 'had',  
'just', 'adopted', 'them', '.', 'Hang', 'it', ',',  
'Rickham', ',', 'with', 'that', 'face', 'watching', 'me', 'I',  
'couldn', '"', 't', 'do', 'another', 'stroke', '.', 'The', 'plain',  
'truth', 'was', ',', 'I', 'didn', '"', 't', 'know', 'where', 'to',  
'put', 'it', '--', 'I', 'had', 'never', 'known', ',',  
'Only', ',', 'with', 'my', 'sitters', 'and', 'my', 'public', ',', 'a',  
'showy', 'splash', 'of', 'colour', 'covered', 'up', 'the', 'fact',  
'--', 'I', 'just', 'threw', 'paint', 'into', 'their', 'faces', '.',  
'.', 'Well', ',', 'paint', 'was', 'the', 'one', 'medium',  
'those', 'dead', 'eyes', 'could', 'see', 'through', '--', 'see',  
'straight', 'to', 'the', 'tottering', 'foundations', 'underneath',  
'.', 'Don', '"', 't', 'you', 'know', 'how', ',', 'in', 'talking', 'a',  
'foreign', 'language', ',', 'even', 'fluently', ',', 'one', 'says',  
'half', 'the', 'time', 'not', 'what', 'one', 'wants', 'to', 'but',  
'what', 'one', 'can', '?', 'Well', '--', 'that', 'was', 'the', 'way',  
'I', 'painted', ';', 'and', 'as', 'he', 'lay', 'there', 'and',  
'watched', 'me', ',', 'the', 'thing', 'they', 'called', 'my', '"',  
'technique', '"', 'collapsed', 'like', 'a', 'house', 'of', 'cards',  
'.', 'He', 'didn', '"', 't', 'sneer', ',', 'you', 'understand', ',',  
'poor', 'Stroud', '--', 'he', 'just', 'lay', 'there', 'quietly',

'watching', 'and', 'on', 'his', 'lips', 'through', 'the', 'gray', 'beard', 'I', 'seemed', 'to', 'hear', 'the', 'question', 'Are', 'you', 'sure', 'you', 'know', 'where', 'you', 're', 'coming', 'out', '?', 'If', 'I', 'could', 'have', 'painted', 'that', 'face', 'with', 'that', 'question', 'on', 'it', 'I', 'should', 'have', 'done', 'a', 'great', 'thing', 'The', 'next', 'greatest', 'thing', 'was', 'to', 'see', 'that', 'I', 'couldn', 't', 'and', 'that', 'grace', 'was', 'given', 'me', 'But', 'oh', 'at', 'that', 'minute', 'Rickham', 'was', 'there', 'anything', 'on', 'earth', 'I', 'wouldn', 't', 'have', 'given', 'to', 'have', 'Stroud', 'alive', 'before', 'me', 'and', 'to', 'hear', 'him', 'say', 'It', 's', 'not', 'too', 'late', 'I', 'll', 'show', 'you', 'how', '?', 'It', 'was', 'too', 'late', 'it', 'would', 'have', 'been', 'even', 'if', 'he', 'd', 'been', 'alive', 'I', 'packed', 'up', 'my', 'traps', 'and', 'went', 'down', 'and', 'told', 'Mrs', 'Stroud', 'Of', 'course', 'I', 'didn', 't', 'tell', 'her', 'that', 'it', 'would', 'have', 'been', 'Greek', 'to', 'her', 'I', 'simply', 'said', 'I', 'couldn', 't', 'paint', 'him', 'that', 'I', 'was', 'too', 'moved', 'She', 'rather', 'liked', 'the', 'idea', 'she', 's', 'so', 'romantic', 'It', 'was', 'that', 'that', 'made', 'her', 'give', 'me', 'the', 'donkey', 'But', 'she', 'was', 'terribly', 'upset', 'at', 'not', 'getting', 'the', 'portrait', 'she', 'did', 'so', 'want', 'him', 'done', 'by', 'some', 'one', 'showy', 'At', 'first', 'I', 'was', 'afraid', 'she', 'wouldn', 't', 'let', 'me', 'off', 'and', 'at', 'my', 'wits', 'end', 'I', 'suggested', 'Grindle', 'Yes', 'it', 'was', 'I', 'who', 'started', 'Grindle', 'I', 'told', 'Mrs', 'Stroud', 'he', 'was', 'the', 'coming', 'man', 'and', 'she', 'told', 'somebody', 'else', 'and', 'so', 'it', 'got', 'to', 'be', 'true', 'And', 'he', 'painted', 'Stroud', 'without', 'wincing', 'and', 'she', 'hung', 'the', 'picture', 'among', 'her', 'husband', 's', 'things', 'He', 'flung', 'himself', 'down', 'in', 'the', 'arm-chair', 'near', 'mine', 'laid', 'back', 'his', 'head', 'and', 'clasping', 'his', 'arms', 'beneath', 'it', 'looked', 'up', 'at', 'the', 'picture', 'above', 'the', 'chimney-piece', 'I', 'like', 'to', 'fancy', 'that', 'Stroud', 'himself', 'would', 'have', 'given', 'it', 'to', 'me', 'if', 'he', 'd', 'been', 'able', 'to', 'say', 'what', 'he', 'thought', 'that', 'day', 'And', 'in', 'answer', 'to', 'a', 'question', 'I', 'put', 'half-mechanically', 'Begin', 'again', '?', 'he', 'flashed', 'out', 'When', 'the', 'one', 'thing', 'that', 'brings', 'me', 'anywhere', 'near', 'him', 'is', 'that', 'I', 'knew', 'enough', 'to', 'leave', 'off', '?', 'He', 'stood', 'up', 'and', 'laid', 'his', 'hand', 'on', 'my', 'shoulder', 'with', 'a', 'laugh', 'Only', 'the', 'irony', 'of', 'it', 'is', 'that', 'I', 'am', 'still', 'painting', 'since', 'Grindle', 's', 'doing', 'it', 'for', 'me', '!

```
'The', 'Strouds', 'stand', 'alone', ',', 'and', 'happen', 'once',  
'--', 'but', 'there', '"', 's', 'no', 'exterminating', 'our', 'kind',  
'of', 'art', '.', '']
```

Fetching first 30 words

```
print(preprocessed[:30])  
  
['I', 'HAD', 'always', 'thought', 'Jack', 'Gisburn', 'rather', 'a',  
'cheap', 'genius', '--', 'though', 'a', 'good', 'fellow', 'enough',  
--, 'so', 'it', 'was', 'no', 'great', 'surprise', 'to', 'me', 'to',  
'hear', 'that', ',', 'in']
```

Counting the number of words

```
all_words = sorted(set(preprocessed))  
vocab_size = len(all_words)  
print(vocab_size)
```

1130

Creating a dictionary with tokens (words) and their corresponding token ids

```
vocab = {token:integer for integer,token in enumerate(all_words)}  
for i, item in enumerate(vocab.items()):  
    print(item)  
    if i >= 50:  
        break
```

```
('!', 0)  
('"', 1)  
('\"', 2)  
('(', 3)  
(')', 4)  
(',', 5)  
('--', 6)  
('.', 7)  
(':', 8)  
('; ', 9)  
('?', 10)  
('A', 11)  
('Ah', 12)  
('Among', 13)  
('And', 14)  
('Are', 15)  
('Arret', 16)  
('As', 17)  
('At', 18)  
('Be', 19)  
('Begin', 20)  
('Burlington', 21)
```

```
('But', 22)
('By', 23)
('Carlo', 24)
('Chicago', 25)
('Claude', 26)
('Come', 27)
('Croft', 28)
('Destroyed', 29)
('Devonshire', 30)
('Don', 31)
('Dubarry', 32)
('Emperors', 33)
('Florence', 34)
('For', 35)
('Gallery', 36)
('Gideon', 37)
('Gisburn', 38)
('Gisburns', 39)
('Grafton', 40)
('Greek', 41)
('Grindle', 42)
('Grindles', 43)
('HAD', 44)
('Had', 45)
('Hang', 46)
('Has', 47)
('He', 48)
('Her', 49)
('Hermia', 50)
```

Below are the functions to encode into tokens and token ids then decode back to normal text

```
class SimpleTokenizerV1:
    def __init__(self, vocab):
        self.str_to_int = vocab
        self.int_to_str = {i:s for s,i in vocab.items()}
    def encode(self, text):
        preprocessed = re.split(r'([,._!"()\']|--|\s)', text)
        preprocessed = [
            item.strip() for item in preprocessed if item.strip()
        ]
        ids = [self.str_to_int[s] for s in preprocessed]
        return ids
    def decode(self, ids):
        text = " ".join([self.int_to_str[i] for i in ids])
        text = re.sub(r'\s+([,._!"()\'])', r'\1', text)
        return text
```

We use the above encode function on random text

```
tokenizer = SimpleTokenizerV1(vocab)
text = """It's the last he painted, you know,"
Mrs. Gisburn said with pardonable pride."""
ids = tokenizer.encode(text)
print(ids)

[1, 56, 2, 850, 988, 602, 533, 746, 5, 1126, 596, 5, 1, 67, 7, 38,
851, 1108, 754, 793, 7]
```

Next, we use the decode function to decode based on token ids generated in the previous step

```
print(tokenizer.decode(ids))

" It' s the last he painted, you know," Mrs. Gisburn said with
pardonable pride.
```

The code below checks if the above process works with unknown words -- unfortunately, it doesn't

```
text = "Hello, do you like tea?"
print(tokenizer.encode(text))

-----
-----
KeyError                                Traceback (most recent call
last)
Cell In[44], line 2
      1 text = "Hello, do you like tea?"
----> 2 print(tokenizer.encode(text))

Cell In[38], line 10, in SimpleTokenizerV1.encode(self, text)
      6 preprocessed = re.split(r'([,._!"()\'|]|--|\s)', text)
      7 preprocessed = [
      8     item.strip() for item in preprocessed if item.strip()
      9 ]
--> 10 ids = [self.str_to_int[s] for s in preprocessed]
     11 return ids

KeyError: 'Hello'
```

Below code handles unknown words with <|unk|> token and end of a segment with <|endoftext|> token

```
all_tokens = sorted(list(set(preprocessed)))
all_tokens.extend(["<|endoftext|>", "<|unk|>"])
vocab = {token:integer for integer,token in enumerate(all_tokens)}
print(len(vocab.items()))
```

1132

Fetch last 5 entries with token and token ids

```
for i, item in enumerate(list(vocab.items())[-5:]):
    print(item)

('younger', 1127)
('your', 1128)
('yourself', 1129)
('<|endoftext|>', 1130)
('<|unk|>', 1131)
```

We incorporate the <|unk|> token and <|endoftext|> token into the encode and decode functions

```
class SimpleTokenizerV2:
    def __init__(self, vocab):
        self.str_to_int = vocab
        self.int_to_str = { i:s for s,i in vocab.items()}
    def encode(self, text):
        preprocessed = re.split(r'([,.;?!"()\\']|--|\s)', text)
        preprocessed = [
            item.strip() for item in preprocessed if item.strip()
        ]
        preprocessed = [item if item in self.str_to_int
            else "<|unk|>" for item in preprocessed]
        ids = [self.str_to_int[s] for s in preprocessed]
        return ids
    def decode(self, ids):
        text = " ".join([self.int_to_str[i] for i in ids])
        text = re.sub(r'\s+([,.;?!"()\\'])', r'\1', text)
        return text
```

Application of encode and decode

```
tokenizer = SimpleTokenizerV2(vocab)
print(tokenizer.encode(text))

[1131, 5, 355, 1126, 628, 975, 10, 1130, 55, 988, 956, 984, 722, 988,
1131, 7]

print(tokenizer.decode(tokenizer.encode(text)))

<|unk|>, do you like tea? <|endoftext|> In the sunlit terraces of the
<|unk|>.
```