

```

import csv
from collections import defaultdict

# Step 1: Read the CSV file and store unique non-stop words in a set
unique_non_stop_words = set()

file_path = 'all_ns_words.csv' # Input file path

with open(file_path, 'r', encoding='utf-8') as csv_file:
    reader = csv.reader(csv_file)
    header = next(reader) # Read the header row

    for row in reader:
        word = row[0]
        unique_non_stop_words.add(word)

# Step 2: Calculate the number of unique non-stop words
num_unique_non_stop_words = len(unique_non_stop_words)

print("Number of Unique Non-Stop Words:", num_unique_non_stop_words)

```

Number of Unique Non-Stop Words: 3445

```

import csv

# Step 1: Read the CSV file and store word probabilities in a dictionary
word_probabilities = {}

file_path = 'all_ns_words.csv' # Input file path

with open(file_path, 'r', encoding='utf-8') as csv_file:
    reader = csv.reader(csv_file)
    header = next(reader) # Read the header row

    for row in reader:
        word = row[0]
        probability = float(row[2])
        word_probabilities[word] = probability

# Step 2: Find the least probable word(s)
min_probability = min(word_probabilities.values())
least_probable_words = [word for word, probability in
word_probabilities.items() if probability == min_probability]

print("Least Probable Word(s):", least_probable_words)
print("Probability:", min_probability)

```

Least Probable Word(s): ['1500', '15961650', '16461716', '16851753', '171176', '17241804', '17701831', '1912', '1a', '1b', '1d', '1e2', '1e3', '1e4', '1e5', '1e6', '1f', '1f1', '1f2', '1f4', '1f5', '1f6',

'20', '2001', '2004', '2019', '30', '50', '5000', '501c3', '5827',
'5961887', '60', '646221541', '801', '809', '84116', 'abide',
'abroad', 'abstain', 'abstracted', 'abstractions', 'abstractly',
'absurdities', 'acceptance', 'accepts', 'accessed', 'accessible',
'accident', 'accidental', 'accidents', 'accompanied', 'accord',
'accounting', 'accurate', 'accusative', 'achieve', 'achieves',
'acquaintedusually', 'acquiesce', 'acquiescence', 'acquisition',
'activity', 'acute', 'ad', 'adapts', 'adding', 'address', 'addresses',
'adduced', 'adequacy', 'adequately', 'admirable', 'admittedthat',
'admitthrough', 'admitting', 'admixture', 'advances', 'adversaries',
'advocate', 'advocates', 'advocating', 'aesthetic', 'aether', 'af',
'affairs', 'affections', 'affirmative', 'afforded', 'agent',
'agreeably', 'agreedthe', 'aim', 'aimed', 'akin', 'alien',
'allembracing', 'allimportant', 'allowed', 'allowing', 'allows',
'allround', 'alteration', 'altered', 'alternate', 'alternatives',
'alters', 'ambitious', 'amend', 'amply', 'analogous', 'analyse',
'analyses', 'anatomist', 'animal', 'announcement', 'anticipate',
'anticipating', 'anticipation', 'anybodys', 'appearances',
'appearing', 'appetite', 'application', 'applied', 'apportions',
'apprehendedmust', 'apprehends', 'approaches', 'approaching',
'arduous', 'argued', 'arguedcorrectly', 'argues', 'arguing', 'arose',
'arrange', 'array', 'arrived', 'arriving', 'arrogant', 'artificial',
'ascertained', 'asks', 'assassinated', 'assented', 'assigns',
'assimilate', 'assist', 'assumptions', 'assurance', 'assure',
'assured', 'astray', 'ate', 'atheists', 'atoms', 'attached', 'attack',
'attain', 'attainable', 'attainedmany', 'attested', 'attitude',
'attributing', 'author', 'avoided', 'avoiding', 'awaresay', 'away\
x97you', 'babies', 'backward', 'backwards', 'badit', 'baldness',
'banging', 'bannerman', 'bar', 'barrier', 'beat', 'became',
'becoming', 'bed', 'beg', 'begging', 'beginner', 'begs', 'begun',
'behave', 'behaviour', 'beingas', 'beis', 'beleagured', 'beliefsfor',
'beliefssuch', 'believesto', 'bell', 'besomething', 'bewildering',
'bibliographical', 'bid', 'bigger', 'binary', 'binds', 'bismarcks',
'blank', 'bluegreen', 'bolder', 'bone', 'bored', 'borne', 'bounds',
'bradley', 'brick', 'brief', 'briefest', 'brighter', 'brightness',
'british', 'broken', 'brotherhood', 'build', 'builder', 'buildings',
'calculate', 'calculated', 'calm', 'campbell', 'campsfriends',
'candid', 'cantor', 'card', 'careful', 'carelessly', 'carry',
'catalogue', 'catalogued', 'catching', 'categories', 'centuries',
'century', 'cf', 'chance', 'changeable', 'changed', 'changes',
'charitable', 'charities', 'checks', 'child', 'chimaera', 'choice',
'chosen', 'citizens', 'citizenship', 'clash', 'clashes', 'clean',
'clock', 'closed', 'closer', 'closes', 'clouds', 'coalesce', 'codes',
'cogito', 'cohere', 'coin', 'coins', 'collective', 'colony',
'colourblind', 'combination', 'combine', 'coming', 'commercial',
'commit', 'communicate', 'community', 'comparative', 'compare',
'compared', 'comparing', 'competent', 'compilation', 'complement',
'completed', 'compressed', 'conceivable', 'conceivably', 'conceiving',
'concentrate', 'concluded', 'concourse', 'condemns', 'conduct',

'confess', 'confine', 'confining', 'confirmation', 'confirmed',
'confirms', 'conflict', 'conformity', 'confused', 'confusing',
'confusions', 'confute', 'conjecture', 'connaître', 'connecting',
'connects', 'consent', 'consequences', 'consequential',
'consequently', 'consisted', 'constituting', 'construct',
'constructed', 'constructing', 'constructions', 'constructive',
'contemplated', 'contemplating', 'contemplative', 'contemptuously',
'contended', 'contention', 'contentions', 'continental', 'continued',
'continuing', 'continuous', 'continuously', 'contract', 'contradict',
'contradictory', 'contrast', 'contributed', 'contribution',
'convenient', 'converged', 'conversation', 'convert', 'convincing',
'cool', 'cooperation', 'corporation', 'correctly', 'correlative',
'corresponded', 'corrupt', 'counterparts', 'countless', 'countries',
'cranny', 'creates', 'creation', 'criticisms', 'criticized',
'crudely', 'cs', 'current', 'custom', 'customary', 'cutting',
'damaged', 'danger', 'dataif', 'dates', 'debate', 'debated',
'deceitful', 'deceive', 'december', 'deceptive', 'decides',
'deciding', 'decision', 'declared', 'deducing', 'deductible', 'deeds',
'deeply', 'defeated', 'defensible', 'defined', 'definiteness',
'definitions', 'deletions', 'deliberate', 'delightful', 'delusion',
'demanding', 'demonstrably', 'demonstrative', 'demonstratively',
'denies', 'dentist', 'dependence', 'depending', 'descend',
'describes', 'describing', 'descriptive', 'desertit', 'deserts',
'deserves', 'deserving', 'designate', 'desirability', 'desirable',
'desiring', 'desirous', 'despite', 'destroyed', 'destroys',
'destructive', 'detach', 'determines', 'development', 'developments',
'diary', 'dictionary', 'die', 'different something', 'differingsome',
'diminish', 'diminished', 'diminishes', 'dined', 'dinnertable',
'directions', 'directs', 'disappear', 'disappoint', 'disappointing',
'disclaim', 'disclaimers', 'discontinue', 'discuss', 'disease',
'disengaged', 'disk', 'dismiss', 'dispassionately', 'display',
'displayed', 'disputable', 'disputants', 'dissociated', 'distinctly',
'distinguishes', 'distort', 'distorts', 'distributor', 'diverge',
'diversity', 'divest', 'divide', 'divided',
'divisibility philosophers', 'dogmas', 'dogmatic', 'dogmatism',
'dominion', 'donation', 'donors', 'doubtless', 'downloading', 'drawn',
'dream that', 'dreamtable', 'dried', 'drifted', 'drive', 'drives',
'duly', 'e', 'early', 'easiest', 'east', 'eclipse', 'edition',
'educated', 'educational', 'effected', 'ein', 'elapse', 'elapsed',
'elect', 'elected', 'electric', 'elementary', 'elements',
'else something', 'elsewhere', 'elucidation', 'embark', 'emerged',
'emerges', 'emitted', 'emitting', 'emphasized', 'emphatically',
'empiricists who', 'employee', 'employees', 'employs', 'engagement',
'engendered', 'england', 'english', 'enlarged', 'enlarges',
'enormously', 'enquiry', 'enrich', 'ensuring', 'entails', 'entangled',
'enter', 'entertaining', 'enumerated', 'enumeration', 'enunciated',
'equal', 'equivocation', 'ergo', 'esse', 'estimated', 'etcis',
'etc may', 'etc which', 'eternal', 'eternally', 'euclids', 'europe',
'everyday', 'exact', 'examines', 'exampleraises', 'exceedingly',

'excluding', 'exclusion', 'exemplify', 'existent', 'expend',
'expense', 'experienceas', 'experiencenot', 'explicitly', 'explored',
'exploring', 'exponents', 'exporting', 'expressions', 'extend',
'extends', 'extensive', 'extrinsic', 'fabric', 'faceit', 'facility',
'factsinfinite', 'faculty', 'failing', 'fainter', 'faintness',
'fallacies', 'fallaciouswhich', 'fallacy', 'fallible', 'falls',
'falsified', 'family', 'fancy', 'farther', 'fast', 'fatal',
'favouritism', 'fear', 'features', 'fed', 'feeds', 'fetters',
'feverish', 'file', 'files', 'financial', 'fine', 'fitness', 'fix',
'fixed', 'flash', 'foes', 'followed', 'foolish', 'football', 'force',
'forces', 'forget', 'forgetting', 'forgo', 'forgotten', 'forks',
'formal', 'formally', 'formats', 'fortress', 'fortuitous', 'fortune',
'forty', 'forwards', 'fostered', 'fourth', 'fragment', 'fragments',
'framework', 'friend', 'frightened', 'front', 'fulfilment', 'fulfils',
'g', 'gain', 'garrison', 'generality', 'genuinelyempirical',
'geography', 'geology', 'georg', 'george', 'getting', 'ghoststories',
'gilbert', 'glance', 'glasses', 'globe', 'goals', 'gone', 'goodwill',
'gordon', 'gradations', 'grain', 'grapple', 'grasp', 'gratefully',
'gratuitous', 'gravely', 'greeks', 'greens', 'greenyblue',
'groundless', 'group', 'grow', 'growing', 'guarded', 'guide',
'gutenbergs', 'gutenberg\x99\x92s', 'habits', 'habitual',
'habitually', 'hairsplitting', 'halves', 'handbooks', 'happiness',
'hardnesses', 'harm', 'harmless', 'harmony', 'hart', 'hates',
'hatred', 'hatreds', 'heartbeats', 'heaven', 'heavens', 'hegels',
'helpful', 'henry', 'heretofore', 'hesitatingly', 'hides', 'high',
'higher', 'hills', 'hinder', 'historian', 'historic', 'historically',
'home', 'honour', 'hoofs', 'hoped', 'hopeless', 'horrors', 'host',
'hot', 'howeverwhich', 'humemaintained', 'humes', 'hundred',
'hundreds', 'hypertext', 'ideasie', 'identification', 'identity',
'ides', 'ignorance', 'ignorant', 'ii', 'iii', 'illumination',
'illusoriness', 'illusory', 'illustrated', 'illustrations',
'immanuel', 'immutable', 'impaired', 'impairs', 'impartial',
'impartially', 'impenetrable', 'impersonal', 'impose', 'impression',
'imprisoned', 'inaccurate', 'incidental', 'inclination', 'includes',
'incompatible', 'inconsistencies', 'inconsistency', 'increase',
'increasing', 'incredulous', 'indefinitely', 'indemnify', 'indemnity',
'independence', 'indestructible', 'indicated', 'indicating',
'indirect', 'inevitable', 'inexplicable', 'infallibility', 'infancy',
'infected', 'inferring', 'infinitesimal', 'infinitem', 'inflected',
'inflections', 'influence', 'infringed', 'infringement',
'inhabitants', 'inhabited', 'inherent', 'innocent', 'inoperative',
'inquire', 'inquired', 'insistent', 'insoluble', 'inspiration',
'instants', 'instincts', 'instrument', 'insubstantial', 'intelligent',
'intend', 'intended', 'intently', 'interact', 'interferes',
'intermediary', 'international', 'interpret', 'interpreted',
'interrelation', 'interrupted', 'introduce', 'introduction',
'intuitively', 'invalidity', 'invented', 'inventing', 'invention',
'inventions', 'investigations', 'invites', 'involvedat', 'involvedin',
'irrefutable', 'irs', 'isolation', 'iti', 'itself1',

'ittheoretically', 'j', 'jealousy', 'joint', 'june', 'justly',
'kantian', 'keener', 'keeps', 'kennen', 'keynes', 'killed', 'kings',
'knits', 'knives', 'knowable', 'knowledgefar', 'knowledgeknowledge',
'königsberg', 'lake', 'lapse', 'larger', 'latin', 'launched', 'lay',
'learnt', 'legally', 'legitimately', 'leibnizmaintained', 'length',
'lessfrom', 'level', 'liberating', 'liberation', 'liberator',
'liberty', 'licensed', 'lies', 'lifelong', 'lifewhich', 'lightwaves',
'limitations', 'limiting', 'linked', 'lips', 'live', 'lived',
'locations', 'locke', 'logicallyso', 'logician', 'longestlived',
'longlived', 'looked', 'loose', 'loosen', 'lot', 'loud', 'lowest',
'luminously', 'lying', 'machinereadable', 'magnifies', 'maintain',
'maintaining', 'majority', 'manager', 'manmade', 'manyprofess',
'march', 'marching', 'marriage', 'match', 'maximum', 'meanings',
'meditations', 'melt', 'mens', 'mention', 'merchantability',
'mercilessly', 'merest', 'merges', 'merits', 'metaphorically',
'metaphysic', 'michael', 'middle', 'middleaged', 'midst', 'miles',
'million', 'mindnot', 'mineralogist', 'minor', 'miracle', 'misery',
'misleadingness', 'mississippi', 'mistake', 'mistakes',
'misunderstanding', 'mitigate', 'mode', 'modification',
'modifications', 'modify', 'momentary', 'monad', 'monadism',
'monadology', 'monism', 'month', 'moore', 'moved', 'movements',
'multiplication', 'multiplicity', 'murray', 'musical', 'mutually',
'mystic', 'mystical', 'mysticism', 'n', 'nation', 'native', 'nearest',
'nearness', 'needs', 'neglecting', 'negligence', 'negligible',
'network', 'newness', 'news', 'newsletter', 'newspapers', 'newtons',
'ninetythree', 'nobody', 'noises', 'nominative', 'nonexistence',
'nonlogical', 'nonprofit', 'nonproprietary', 'nook', 'note',
'noteworthy', 'noticing', 'notifies', 'notions', 'nouns', 'novelist',
'nowadays', 'nowhen', 'nowhere', 'numerous', 'objectin',
'objectionable', 'objective', 'objectsit', 'oblivious', 'obscure',
'observe', 'observing', 'obsolete', 'obstacle', 'obstacles',
'obstinate', 'obtuse', 'obviousness', 'obviousnessthe', 'occupant',
'occupying', 'odd', 'office', 'oftener', 'omission', 'ones',
'oneself', 'one\97the', 'opaque', 'opening', 'operated', 'operative',
'opportunities', 'opportunity', 'opposites', 'orator', 'orderly',
'orders', 'ordnance', 'organized', 'organizing', 'organs', 'origin',
'originally', 'originator', 'outcome', 'outdated', 'outlines',
'outward', 'oval', 'owed', 'painfully', 'palpable', 'papers',
'paperwork', 'paradoxes', 'paradoxical', 'partially', 'partlyand',
'party', 'passage', 'passes', 'pastnor', 'pastnot', 'pausing', 'pay',
'peace', 'peculiar', 'peculiarly', 'percipi', 'perfection',
'performances', 'periodic', 'permanence', 'perpetual', 'persons',
'perspective', 'persuaded', 'persuading', 'perturbed', 'pg', 'pglaf',
'phantasmagoria', 'philosophersor', 'philosophersthat',
'philosophize', 'philosophyfor', 'philosophythe', 'physics',
'physiological', 'physiology', 'pink', 'placeslondon', 'placing',
'plane', 'planet', 'plausibility', 'playing', 'pointed', 'poisonous',
'positively', 'possess', 'possessed', 'possession', 'poverty',
'powerlessness', 'powers', 'practice', 'practised', 'precisely',

'preeminently', 'preface', 'preferable', 'preparing', 'prescribe',
'presence', 'presented', 'presents', 'preservation', 'pressing',
'pressure', 'pressures', 'presumption', 'presupposed', 'pretty',
'prevent', 'previously', 'primitive', 'principlesperhaps',
'proceeded', 'proceeds', 'processes', 'processing', 'producing',
'production', 'products', 'profitable', 'profited', 'profits',
'profitthe', 'profoundest', 'progress', 'progressively',
'prohibition', 'prolegomena', 'prolonged', 'promote', 'promotion',
'pronounce', 'proofread', 'proposed', 'proprietary', 'provision',
'provisionally', 'provisions', 'prudent', 'prussia',
'psychologically', 'psychology', 'punitive', 'pursue', 'pushed',
'puts', 'puzzling', 'questionsand', 'quibbling', 'quickly', 'range',
'ranging', 'rapping', 'rationalist', 'rationalistswho', 'rationality',
'readable', 'readers', 'realization', 'realized', 'realor',
'reappear', 'reasonsin', 'receiving', 'recently', 'recognizes',
'recombines', 'reconcile', 'reconstructed', 'recorded',
'redistribute', 'reduce', 'reducible', 'refined', 'reflecting',
'refund\x94', 'refutable', 'regress', 'regulating', 'rejection',
'relate', 'relational', 'relationin', 'release', 'reliable',
'religion', 'remarkably', 'remarks', 'remedies', 'remote', 'remoter',
'removal', 'removes', 'renamed', 'renounce', 'repeat', 'repeatedly',
'repetition', 'replaced', 'reported', 'reports', 'represent',
'representations', 'representative', 'representing', 'republic',
'request', 'resemblances', 'residue', 'restate', 'resting',
'restricted', 'retain', 'retained', 'retort', 'returning', 'returns',
'revealed', 'revenue', 'reversed', 'revert', 'review', 'revolution',
'rigid', 'rival', 'rob', 'robbing', 'roman', 'roof', 'rooted',
'rotate', 'rotating', 'roundabout', 'rounded', 'rouse', 'roused',
'runs', 'salt', 'samples', 'satisfaction', 'satisfying', 'save',
'savoir', 'scaffold', 'sceptic', 'sceptical', 'sceptics', 'schoolmen',
'sea', 'searches', 'sections', 'secures', 'seldom',
'selfacquaintedwithsensedatum', 'selfinterest', 'selfsubsistent',
'send', 'sending', 'sensationthe', 'sensedatabrown',
'sensedatacolour', 'sensedatafor', 'sensedatawhich', 'sensethoughts',
'sent', 'separated', 'sequel', 'serious', 'serve', 'sets', 'seven',
'seventeenth', 'severe', 'shadow', 'sharing', 'sharp', 'sheets',
'shiny', 'shock', 'short', 'shortly', 'showed', 'significantly',
'simplify', 'sir', 'situated', 'skeleton', 'sleeping', 'slowly',
'smallest', 'smelling', 'smells', 'sober', 'society', 'sold',
'solicitation', 'solid', 'solve', 'solves', 'solving', 'somebody',
'somebodys', 'sophistry', 'soul', 'sour', 'south', 'spacerelations',
'speakthat', 'specially', 'species', 'spectacles', 'speculative',
'speech', 'spoke', 'spoons', 'sprang', 'staff', 'stages', 'started',
'starts', 'state\x92s', 'step', 'stone', 'stored', 'straightforward',
'strangeness', 'stranger', 'strangest', 'strike', 'strikes',
'striking', 'strives', 'striving', 'struck', 'structure',
'struggling', 'struldbugs', 'studies', 'sublime', 'subscribe',
'subsequent', 'subsequently', 'subsumed', 'subsumptions', 'success',
'successively', 'suffering', 'sufficed', 'sufficiently', 'suggest',

'suggestion', 'suggestions', 'sumtotal', 'suns', 'sunset', 'supplied',
'suprasensible', 'surprised', 'surrender', 'surrounded', 'surveys',
'survive', 'surviving', 'suspended', 'swamp', 'sweet', 'swift',
'synonymous', 'systematize', 'systemsimilarly', 'tableand',
'tableare', 'tablecloth', 'tableit', 'tableits', 'tactile', 'tap',
'task', 'tasting', 'taught', 'taxes', 'telegram', 'temper',
'temperaments', 'tempted', 'ten', 'tending', 'tenet', 'term', 'texts',
'texture', 'theletters', 'thence', 'theoretical', 'thereby', 'theses',
'thesis', 'thisthat', 'thoroughly', 'thraldom', 'throw', 'throws',
'thwarting', 'timeless', 'timerelations', 'tincture', 'title',
'tolerably', 'tongue', 'took', 'toothache', 'topics', 'total',
'touches', 'touchleaves', 'trademarkcopyright', 'traditional',
'trains', 'trammels', 'transcribe', 'transcription', 'transfer',
'transform', 'transitory', 'travels', 'treated', 'treating',
'treatment', 'tries', 'trifling', 'trotting', 'trouble', 'troubled',
'troubles', 'trusted', 'trying', 'tuning', 'turning', 'twice', 'type',
'types', 'tyranny', 'unaccustomed', 'unaffected', 'unattainable',
'uncertain', 'unchanged', 'unchanging', 'uncomfortable', 'uncommon',
'unconscious', 'undeniable', 'undeniably', 'underlie', 'underlies',
'understandin', 'understanding', 'undertake', 'undiminished', 'undue',
'uneducated', 'unenforceability', 'unexpectedly', 'unexperienced',
'unimportant', 'unites', 'uniting', 'university', 'unlearn', 'unlink',
'unnecessary1', 'unplausible', 'unprotected', 'unpublished',
'unreflectingly', 'unreflective', 'unsafe', 'unsolicited', 'unsolved',
'unsound', 'unsupported', 'unsuspected', 'unusual', 'unwarrantable',
'unwise', 'upset', 'upstairs', 'urges', 'usage', 'ut', 'v',
'vagueness', 'vaguer', 'vain', 'valleys', 'valueperhaps',
'valuethrough', 'variable', 'variations', 'variety', 'variously',
'varying', 'vast', 'veil', 'verification', 'version', 'viewing',
'viewwhich', 'viii', 'vindicate', 'violates', 'violent', 'virtuous',
'virus', 'voice', 'voices', 'void', 'volume', 'volunteer', 'voyage',
'walk', 'walks', 'walled', 'walls', 'wants', 'warrants', 'waste',
'waterloo', 'wear', 'wearing', 'weather', 'web', 'weight', 'whence',
'whitehead', 'whoever', 'willing', 'window', 'wise', 'wissen',
'wonderful', 'wooden', 'wore', 'worked', 'working', 'worst', 'wrings',
'writers', 'wwwgutenbergorgcontact', 'wwwgutenbergorglicense', 'x',
'xii', 'xiii', 'xiv', 'xv', 'yield', 'york', '\x91asis\x92', '\x93defects\x94', '\x93information', '\x93right', '\x93the']
Probability: 4.699027301348621e-05

Answer 3: The probabilities of non-stop words vary, indicating differences in their likelihood of appearing in the text. The sum of probabilities represents the overall likelihood of encountering non-stop words, and there may be ties for the least probable word(s), indicating rare occurrences. The number of unique non-stop words reflects diversity in the text's vocabulary.

```
import csv
import string

def calculate_combined_probability(sentence, word_probabilities):
    words = sentence.split()
    combined_probability = 1.0
    stopwords = []
    stopwords_file = 'stopwords-en.txt'
    with open(stopwords_file, 'r') as file:
        stopwords = set(line.strip() for line in file)

    for word in words:
        if word.lower() not in stopwords:
            cleaned_word = word.strip(string.punctuation)
            probability = word_probabilities.get(cleaned_word.lower(),
0)
            combined_probability *= probability

    return combined_probability

word_probabilities = {}
all_ns_words = 'all_ns_words.csv'

# Read word probabilities from the CSV file
with open(all_ns_words, 'r', newline='') as csvfile:
    csv_reader = csv.reader(csvfile)
    next(csv_reader)
    for row in csv_reader:
        word_probabilities[row[0]] = float(row[2])

sentence1 = "If a belief is true, it can be deduced it is universal."
sentence2 = "Criticism of knowledge is counter to scientific results."

combined_probability1 = calculate_combined_probability(sentence1,
word_probabilities)
combined_probability2 = calculate_combined_probability(sentence2,
word_probabilities)

if combined_probability1 > combined_probability2:
    print(sentence1 + f" is more likely with a combined probability of
{combined_probability1:.10f}.")
elif combined_probability2 > combined_probability1:
```



```

        print(sentence2 + f" is more likely with a combined probability of
{combined_probability2:.10f}.")
    else:
        print(f"Both the sentences have the same combined probability of
{combined_probability1:.10f}.")

```

```

-----
-----
UnicodeDecodeError                                Traceback (most recent call
last)

```

```

Cell In[3], line 33

```

```

    30 sentence1 = "If a belief is true, it can be deduced it is
universal."
    31 sentence2 = "Criticism of knowledge is counter to scientific
results."

```

```

--> 33 combined_probability1 =
calculate_combined_probability(sentence1, word_probabilities)
    34 combined_probability2 =
calculate_combined_probability(sentence2, word_probabilities)
    36 if combined_probability1 > combined_probability2:

```

```

Cell In[3], line 10, in calculate_combined_probability(sentence,
word_probabilities)

```

```

    8 stopwords_file = 'stopwords-en.txt'
    9 with open(stopwords_file, 'r') as file:
--> 10     stopwords = set(line.strip() for line in file)
    12 for word in words:
    13     if word.lower() not in stopwords:

```

```

Cell In[3], line 10, in <genexpr>(.0)

```

```

    8 stopwords_file = 'stopwords-en.txt'
    9 with open(stopwords_file, 'r') as file:
--> 10     stopwords = set(line.strip() for line in file)
    12 for word in words:
    13     if word.lower() not in stopwords:

```

```

File ~\anaconda3\Lib\encodings\cp1252.py:23, in

```

```

IncrementalDecoder.decode(self, input, final)
    22 def decode(self, input, final=False):
--> 23     return
codecs.charmap_decode(input, self.errors, decoding_table)[0]

```

```

UnicodeDecodeError: 'charmap' codec can't decode byte 0x9d in position
2633: character maps to <undefined>

```

```

# Import necessary libraries

```

```

import csv
import string

```

```

# Function to calculate the combined probability of a sentence

```

```

def calculate_combined_probability(sentence, word_probabilities):
    # Split the sentence into words
    words = sentence.split()
    combined_probability = 1.0

    # Load and preprocess stopwords from a file
    stopwords = []
    stopwords_file = 'stopwords-en.txt'
    with open(stopwords_file, 'r', encoding='utf-8') as file:
        stopwords = set(line.strip() for line in file)

    # Calculate the combined probability by multiplying word
    # probabilities
    for word in words:
        # Check if the word is not a stopword
        if word.lower() not in stopwords:
            cleaned_word = word.strip(string.punctuation) # Remove
            # punctuation
            probability = word_probabilities.get(cleaned_word.lower(),
            0) # Get word probability
            combined_probability *= probability # Multiply the
            # probabilities

    return combined_probability

# Initialize a dictionary to store word probabilities
word_probabilities = {}

# Specify the CSV file containing word probabilities
all_ns_words = 'all_ns_words.csv'

# Read word probabilities from the CSV file and populate the
# dictionary
with open(all_ns_words, 'r', newline='') as csvfile:
    csv_reader = csv.reader(csvfile)
    next(csv_reader) # Skip the header row
    for row in csv_reader:
        word_probabilities[row[0]] = float(row[2]) # Word and its
        # probability

# Define the two sentences for comparison
sentence1 = "If a belief is true, it can be deduced it is universal."
sentence2 = "Criticism of knowledge is counter to scientific results."

# Calculate the combined probabilities for both sentences
combined_probability1 = calculate_combined_probability(sentence1,
word_probabilities)
combined_probability2 = calculate_combined_probability(sentence2,
word_probabilities)

```

```
# Compare the combined probabilities and determine which sentence is more likely
if combined_probability1 > combined_probability2:
    print(sentence1 + f" is more likely with a combined probability of {combined_probability1:.10f}.")
elif combined_probability2 > combined_probability1:
    print(sentence2 + f" is more likely with a combined probability of {combined_probability2:.10f}.")
else:
    print(f"Both sentences have the same combined probability of {combined_probability1:.10f}.")
```

If a belief is true, it can be deduced it is universal. is more likely with a combined probability of 0.0000000000.