**1. Fire Causes Filtering Seeds (13 languages)**

"cigarette", "cigar", "combustible", "welding", "gasoline", "lighter", "malfunction", "wiring", "circuit", "electric", "arson", "human error", "act of god"

**Translated Languages**: Catalan, German, Spanish, French, Italian, Japanese, Dutch, Polish, Portuguese, Swedish, Thai.

"cigarette", "cigar", "combustible", "welding", "gasoline", "lighter",

"misuse", "malfunction", "wiring", "short-circuit", "circuit", "electric",

"deficiency", "collision", "arson", "human error", "act of god",

"cigarret", "cigar", "combustible", "soldadura", "gasolina", "encenedor",

"mal ús", "mal funcionament", "cablejat", "curtcircuit", "circuit", "elèctric",

"deficiència", "col·lisió", "incendi", "error humà", "acte de déu",

"Zigarette", "Zigarre", "brennbar", "Schweißen", "Benzin", "Feuerzeug",

"Missbrauch", "Fehlfunktion", "Verkabelung", "Kurzschluss", "Stromkreis", "elektrisch",

"Mangel", "Kollision", "Brandstiftung", "menschliches Versagen", "höhere Gewalt",

"cigarrillo", "cigarro", "combustible", "soldadura", "gasolina", "encendedor",

"mal uso", "mal funcionamiento", "cableado", "cortocircuito", "circuito", "eléctrico",

"deficiencia", "colisión", "incendio provocado", "error humano", "acto de Dios",

"cigarette", "cigare", "combustible", "soudure", "essence", "briquet",

"mauvaise utilisation", "dysfonctionnement", "câblage", "court-circuit", "circuit", "électrique",

"déficience", "collision", "incendie criminel", "erreur humaine", "cas de dieu",

"sigaretta", "sigaro", "combustibile", "saldatura", "benzina", "accendino",

"uso improprio", "malfunzionamento", "cablaggio", "cortocircuito", "circuito", "elettrico",

"difetto", "collisione", "incendio doloso", "errore umano", "atto di dio",

"タバコ", "葉巻", "可燃性", "溶接", "ガソリン", "ライター",

"誤用", "誤動作", "配線", "短絡", "回路", "電気",

"欠乏", "衝突", "放火", "ヒューマンエラー", "神の行",

"sigaret", "sigaar", "brandbaar", "lassen", "benzine", "aansteker",

"misbruik", "storing", "bedrading", "kortsluiting", "circuit", "elektrisch",

"tekort", "botsing", "brandstichting", "menselijke fout", "daad van god",

"papieros", "cygaro", "łatwopalny", "spawalniczy", "benzyna", "zapalniczka",

"niewłaściwe użycie", "awaria", "okablowanie", "zwarcie", "obwód", "elektryczny",

"niedobór", "kolizja", "podpalenie", "błąd ludzki", "akt Boży",

"cigarro", "charuto", "combustível", "soldagem", "gasolina", "isqueiro",

"mau uso", "mau funcionamento", "fiação", "curto-circuito", "circuito", "elétrica",

"deficiência", "colisão", "incêndio criminoso", "erro humano", "ato de deus",

"cigarett", "cigarr", "brännbar", "svetsning", "bensin", "tändare",

"missbruk", "funktionsfel", "ledningar", "kortslutning", "krets", "elektrisk",

"brist", "kollision", "brandstift", "mänskligt fel", "handling av gud",

"บุหรี่", "ซิการ์", "ติดไฟได้", "การเชื่อม", "น้ำมันเบนซิน", "ไฟแช็ก",

"การใช้ผิดปกติ", "ความผิดปกติ", "การเดินสายไฟ", "ลัดวงจร", "วงจร", "ไฟฟ้า",

"การขาด", "การชนกัน", "การลอบวางเพลิง", "ความผิดพลาดของมนุษย์", "การกระทำของพระเจ้า"

**2. Filter Out Texts**

**Remove Notre Dame University related texts**

'coach', 'fans', 'athlete', 'athletes', 'teammates', 'player', 'faculty', 'professor', 'stadium', 'game', 'league', 'university', 'unversities', 'champion', 'school', 'student council', 'scholarship', 'student-led', 'ticket pricing', 'college', 'hunchback', 'hunch-back', 'semifinals', 'semi', 'finalist', '1st half', 'sports', 'championship', 'tournament', 'baseball', 'hockey', 'scorer', 'volleyball', 'golf', 'basketball', 'football', 'run lead', 'runner', 'soccer', 'softball', 'bowling', 'good team', 'game-winning', 'halftime', 'latter scoring', 'extracurricular', 'miami', 'michigan', 'ohio', 'minnesota', 'texas', 'florida', 'iowa', 'wisconsin', 'penn state', 'louis', 'maryland', 'at harvard', 'denver', 'washington', 'acc final', 'acc grind', 'trails', 'track and field', 'notre dame camp', 'offenses', 'volunteers-americorps', 'cactus shadows', 'practice for cupcake', 'regular season', 'superbowl', 'physical activity', 'notre dame official', 'bts bubar', 'four-star', 'race studies course', 'notre dame prep', '#acc', '#elearning', 'shooter', 'student body president', 'call him notre dame', 'notre dame hat', 'type of atmosphere', 'exatili', 'notre-dame gargoyle', 'students at notre dame', 'cheshire boys', 'support from big brands', 'notre dame and lidar', 'sixth-overall pick', 'lillards', 'notre dame taught me', 'real boy', 'transfer portal', 'father fournier', 'the defense', 'bellway homes', 'lead against', 'miguel delgado', 'organ recitals', 'activists demand','chicagoland', 'chillin at the notre dame', 'limoges boxes', 'students registered', 'notre dame family', 'non-conference schedule', 'brian kelly', 'dj harvey', 'russell westbrook', 'committed to notre dame', 'conference at notre dame', 'duke', 'timeout', 'ties it', 'tailgate', 'georgia', 'look out time','scores', 'hunch back', 'irish', 'dallas', 'face-off', 'go to notre', 'from notre dame'

**Remove “false” fire causes texts**

"Black church", "Saint Sulpice", "Al-Aqsa Mosque", "Odebrecht construction", "Korea", "Spain", "African history", "bodegas", "envigado", "Christian symbols", "black people", "brandstift", "Christian church", "gargoyle", "New York", "another church", "multiple churches", "after notre", "Puerto Rico", "Boise", "quintessence", "Electric chair", "Bollywood", "misbruik", "Orsay", "Shep Smith", "Katie Hopkins", "Seine", "insurance company", "Mother Earth", "exceptional show", "Detroit", "Shepard Smith", "ShepSmith", "kerosene", "American Church", "electric tower", "spiritual arson", "the USA", "falling asleep", "Quasimodo", "rebuild civilization", "miracle of Notre Dame", "humans and animals", "Patrick's Cathedral" "quintessence", "Electric chair", "Bollywood", "EICR", "Katie Hopkins", "Katie Hopkins", "Seine", "NYPD", "arson boy", "Catholic temples ", "Al-Aqsa mosque", "twenty cigarettes", "human-machine", " water or electricity", "electricity or computers", "public urinals", "before electricity", "Patrick’s Cathedral", "in NYC", "hipster fashion", "review the electrical", "Mont San Michel", "church St Sulpice", "St. Patrick", "fire alert in Notre Dame", "use electric elements", "Electrical installations", "Johann Vexo", "stained glass", "Wehrmacht", "electric yard", "cathedrals of Normandy", "goldhighlighter", "a bit lighter", "electric fishing", "Costa Rica", "1986", "Jerusalem", "Disney movies", "stone and wood", "a few other cathedrals", "Eiffel Tower "

**3. Lexicon-based Library**

**Negation Words / Phrases**

"no", "non", "not", "nope", "isn't", "aren't", "wasn't", "weren't", "haven't", "hasn't", "hadn't", "don't", "doesn't", "didn't", "can't", "cannot", "couldn't", "mightn't", "mayn't", "mustn't", "impossible", "shan't", "shouldn't", "won't", "wouldn't", "daren't", "doubt", "needn't", "oughtn't", "unlikely", "nothing", "shocked", "imaginary", "dismiss", "ironic", "doubted", "deny", "denies", "denied", "without", "absence", "declined", "negative", "nobody", "cautious", "strange", "suspicious", "coverup", "rejected", "surprised", "dismissed", "disagree", "exclude", "excluded", "ashamed", "conspiracy", "reject", "fake", "rumor", "upgraded", "replaced", "revised", "restored", "updated", " overhauled", "surprised", "incorrect", "incorrectly", "dispute", "disputed", "pretending"

"rule out", "ruled out", "ruling out", "hard to believe", "political decision", "political statement", "rushed to say", "too soon", "quick to say", "smells fishy", "in spite", "so quick", "ass kicked", "to blame", "be ashamed", "awfully strange"

**Contrast Conjunction Words / Phrases**

"but", "however", "nevertheless", "ironic", "cautious", "deny", "denies", "denied", "strange", "suspicious", "coverup", "nobody", "disagree", "impossible", "though", "stupid", "stupidity", "lie", "lied", "lying", "ridiculous", "fake", "doubt", "doubted", "while", "gullible", "hiding", "hide", "unacceptable", "bullshit", "misleading"

"hard to believe", "political decision","i would really", "difficult to accept", "rushed to say", "too quick", "too soon", "quick to say", "smells fishy", "as long as", "believe different", "even while", "brain worms", "we're not gonna", "all fail", "so quick", "calling bs", "need answers", "as long as", "never be known", "to blame", "too effective to", "doesn’t work", "short circuit in thinking", "🙄"

"awfully strange", "don't believe", "doesn't burn", "can't trust", "gullible idiot"

**Uncertain Words / Phrases**

"don't know", "idk", "hope", "pray", "wish", "hoping", 'investigate whether', "looking into whether"

**Tag Questions**

"don't they", "didn't they", "don't we", "didn't we", "won't they", "won't we", "don't you", "shouldn't they", "shouldn't we", "aren't we", "aren't they", "isn't it", "wasn't it"