

## HW2 Word Embeddings Output:

1. `test_words = ["king", "computer", "Paris", "dog", "doctor"]`

### Most similar words to 'king' ###

CBOW: [('throne', 0.6799127459526062), ('confessor', 0.6648404598236084), ('ruler', 0.6535142660140991), ('abdication', 0.650391161441803), ('anjou', 0.6461496949195862)]

Skip-gram: [('atheling', 0.7576407790184021), ('harefoot', 0.7426010966300964), ('peaceable', 0.7320787310600281), ('unready', 0.7210803627967834), ('underking', 0.7164149284362793)]

GloVe: [('prince', 0.7682328820228577), ('queen', 0.7507690787315369), ('son', 0.7020888328552246), ('brother', 0.6985775232315063), ('monarch', 0.6977890729904175)]

FastText: [('king-', 0.7838029861450195), ('boy-king', 0.7704817652702332), ('queen', 0.7704246640205383), ('prince', 0.7700967192649841), ('kings', 0.7668929696083069)]

### Most similar words to 'computer' ###

CBOW: [('programmer', 0.7322611808776855), ('hardware', 0.7080273032188416), ('graphical', 0.6862488389015198), ('application', 0.6850818395614624), ('processor', 0.6756504774093628)]

Skip-gram: [('hardware', 0.7725837826728821), ('bios', 0.7396647930145264), ('technology', 0.7336676120758057), ('emulator', 0.7221037745475769), ('programmer', 0.7193736433982849)]

GloVe: [('computers', 0.8751984238624573), ('software', 0.8373122215270996), ('technology', 0.7642159461975098), ('pc', 0.7366448640823364), ('hardware', 0.7290390729904175)]

FastText: [('computers', 0.8472651243209839), ('non-computer', 0.8156518936157227), ('mini-computer', 0.8099768757820129), ('micro-computer', 0.7727647423744202), ('super-computer', 0.762492299079895)]

### Most similar words to 'Paris' ###

CBOW: Word not in vocabulary

Skip-gram: Word not in vocabulary

GloVe: Word not in vocabulary

FastText: [('Paris-', 0.7925320863723755), ('France', 0.7642844915390015), ('Parie', 0.744894802570343), ('Paris-Lyon', 0.7436500787734985), ('Paris-Orly', 0.7416573762893677)]

### Most similar words to 'dog' ###

CBOW: [('cowardly', 0.686860203742981), ('poodle', 0.6599791049957275), ('cat', 0.6488620042800903), ('hound', 0.6484981179237366), ('breed', 0.6464619636535645)]

Skip-gram: [('poodle', 0.7822871208190918), ('breed', 0.764072060585022), ('basenji', 0.7357589602470398), ('briard', 0.7176538705825806), ('herding', 0.7163261771202087)]

GloVe: [('cat', 0.8798074126243591), ('dogs', 0.8344309329986572), ('pet', 0.7449564337730408), ('puppy', 0.723637580871582), ('horse', 0.7109653949737549)]

FastText: [('dogs', 0.8457403182983398), ('dogs-', 0.8105224967002869), ('puppy', 0.8071781396865845), ('dog-', 0.776581346988678), ('dog.', 0.7731125354766846)]

### Most similar words to 'doctor' ###

CBOW: [('torchwood', 0.6158494353294373), ('nurse', 0.5797235369682312), ('dumbledore', 0.5501664280891418), ('counselor', 0.5422387719154358), ('rory', 0.5356088280677795)]

Skip-gram: [('medical', 0.659666121006012), ('torchwood', 0.655402660369873), ('osteopathic', 0.6504128575325012), ('pharmacist', 0.6383504867553711), ('nurse', 0.6383141279220581)]

GloVe: [('physician', 0.7673240303993225), ('nurse', 0.75215083360672), ('dr.', 0.7175194025039673), ('doctors', 0.7080884575843811), ('patient', 0.7074184417724609)]

FastText: [('physician', 0.8320724368095398), ('doctor-', 0.7990509271621704), ('non-doctor', 0.7794222831726074), ('doctors', 0.7743481993675232), ('surgeon', 0.7740318775177002)]

2. test\_words = ["king", "computer", "Paris", "dog", "doctor"]

### Most similar words to 'ceo' ###

CBOW: Word not in vocabulary

Skip-gram: Word not in vocabulary

GloVe: [('executive', 0.8861949443817139), ('chairman', 0.7696428298950195), ('chief', 0.7450050115585327), ('cfo', 0.7056118249893188), ('vice', 0.687246561050415)]

FastText: [('Ceo', 0.6498435139656067), ('ceos', 0.6490616202354431), ('C.E.O', 0.6277341842651367), ('sr.', 0.582674503326416), ('CEO', 0.5633218884468079)]

### Most similar words to 'investment' ###

CBOW: [('banking', 0.8533105850219727), ('asset', 0.8203847408294678), ('equity', 0.7892601490020752), ('financial', 0.779140830039978), ('shareholder', 0.7515732049942017)]

Skip-gram: [('banking', 0.8130596876144409), ('brokerage', 0.8064098358154297), ('financial', 0.8037262558937073), ('asset', 0.7779596447944641), ('investor', 0.7753849029541016)]

GloVe: [('investments', 0.8891236782073975), ('fund', 0.7843336462974548), ('asset', 0.7737513780593872), ('financial', 0.7671132683753967), ('firms', 0.7522855997085571)]

FastText: [('investments', 0.8778875470161438), ('re-investment', 0.8094879984855652), ('pre-investment', 0.7942810654640198), ('non-investment', 0.7913498878479004), ('investment-', 0.7627221345901489)]

### Most similar words to 'profit' ###

CBOW: [('invest', 0.8033004403114319), ('payment', 0.7821806073188782), ('surplus', 0.7417104244232178), ('profitable', 0.7361223697662354), ('asset', 0.7268063426017761)]

Skip-gram: [('money', 0.7367629408836365), ('depreciation', 0.7344164848327637), ('taxable', 0.7306875586509705), ('investment', 0.7249315977096558), ('surplus', 0.7204799652099609)]

GloVe: [('profits', 0.9010799527168274), ('earnings', 0.8935472965240479), ('net', 0.811954140663147), ('revenue', 0.8099523186683655), ('sales', 0.7789346575737)]

FastText: [('profits', 0.8384164571762085), ('profitting', 0.7345399260520935), ('profit-', 0.7341814637184143), ('profit.', 0.7299884557723999), ('-profit', 0.7278085947036743)]

### Most similar words to 'startup' ###

CBOW: Word not in vocabulary

Skip-gram: Word not in vocabulary

GloVe: [('start-up', 0.7904508113861084), ('startups', 0.6792274713516235), ('software', 0.6210920214653015), ('incubator', 0.6055338978767395), ('silicon', 0.5928449034690857)]

FastText: [('start-up', 0.9111342430114746), ('startups', 0.7742005586624146), ('Startup', 0.7368545532226562), ('start-ups', 0.7312785387039185), ('Start-up', 0.6989713311195374)]

### Most similar words to 'market' ###

CBOW: [('wholesale', 0.6635810136795044), ('retail', 0.6628846526145935), ('capitalization', 0.6573325991630554), ('stock', 0.6249728798866272), ('demand', 0.6038702726364136)]

Skip-gram: [('wholesale', 0.7177847027778625), ('retail', 0.6948996782302856), ('rasen', 0.6776431798934937), ('capitalization', 0.6704809665679932), ('earner', 0.6535069346427917)]

GloVe: [('markets', 0.8939464092254639), ('stock', 0.7991609573364258), ('prices', 0.7834224700927734), ('stocks', 0.7668591737747192), ('investors', 0.7563610076904297)]

FastText: [('markets', 0.8422666788101196), ('marketplace', 0.8258434534072876), ('home-market', 0.7825994491577148), ('market--and', 0.7811020612716675), ('e-market', 0.7780057787895203)]

### Most similar words to 'enterprise' ###

CBOW: [('flagship', 0.6412880420684814), ('corporation', 0.6152734160423279), ('launch', 0.6135908365249634), ('subsidiary', 0.6045265197753906), ('venture', 0.5967146754264832)]

Skip-gram: [('sife', 0.6404266357421875), ('business', 0.5860429406166077), ('flagship', 0.5844599604606628), ('intrepid', 0.5749533772468567), ('technica', 0.5665135383605957)]

GloVe: [('business', 0.7276235222816467), ('development', 0.6607261896133423), ('management', 0.6600956320762634), ('innovation', 0.649596631526947), ('corporation', 0.6401593089103699)]

FastText: [('enterprises', 0.7827980518341064), ('business', 0.7792075276374817), ('enterprise-scale', 0.7569987773895264), ('enterprise-', 0.7398455739021301), ('enterprise-based', 0.7195351123809814)]

### ### Most similar words to 'entrepreneur' ###

CBOW: [('philanthropist', 0.8465648889541626), ('businessman', 0.8426406383514404), ('billionaire', 0.8102473020553589), ('investor', 0.8016777634620667), ('financier', 0.781832754611969)]

Skip-gram: [('businessman', 0.8547983169555664), ('philanthropist', 0.818892776966095), ('marketer', 0.8079444766044617), ('investor', 0.7843177318572998), ('founder', 0.7609630227088928)]

GloVe: [('businessman', 0.7703410983085632), ('philanthropist', 0.7479178309440613), ('developer', 0.701310932636261), ('billionaire', 0.6780495643615723), ('inventor', 0.6725963950157166)]

FastText: [('businessman', 0.8263760209083557), ('businessperson', 0.7830173969268799), ('entrepreneur', 0.7568551301956177), ('entrepreneurs', 0.7432577013969421), ('businesswoman', 0.7292370200157166)]

### ### Most similar words to 'revenue' ###

CBOW: [('expenditure', 0.7359748482704163), ('tax', 0.7172943949699402), ('asset', 0.7007876038551331), ('taxation', 0.696434736251831), ('excise', 0.6817057132720947)]

Skip-gram: [('tax', 0.7200298309326172), ('expenditure', 0.6953243017196655), ('taxpayer', 0.6922948360443115), ('insurer', 0.6859888434410095), ('crore', 0.6843281388282776)]

GloVe: [('revenues', 0.9188570976257324), ('profits', 0.8107035160064697), ('profit', 0.8099522590637207), ('sales', 0.7821041345596313), ('net', 0.7813507318496704)]

FastText: [('revenues', 0.8840138912200928), ('revenue-', 0.8074648976325989), ('revenue.', 0.7637682557106018), ('non-revenue', 0.7369114756584167), ('tax-revenue', 0.729134738445282)]

### 3. 5 variety

Query 1: 'invest' - 'risk' + 'reward'

{'CBOW': [('pay', 0.6269975304603577)], 'Skip-gram': [('buy', 0.5625247955322266)], 'GloVe': [('invested', 0.637974202632904)], 'FastText': [('reinvest', 0.6238330006599426)]}

-----

Query 2: 'technology' - 'old' + 'new'

{'CBOW': [('innovation', 0.5668476223945618)], 'Skip-gram': [('playability', 0.5604642033576965)], 'GloVe': [('technologies', 0.7700199484825134)], 'FastText': [('technologies', 0.7342677712440491)]}

-----

Query 3: 'student' - 'quit' + 'graduate'

{'CBOW': 'Words not in vocabulary', 'Skip-gram': 'Words not in vocabulary', 'GloVe': [('computer', 0.6472473740577698)], 'FastText': [('undergraduate', 0.631050705909729)]}

-----

Query 4: 'profit' - 'loss' + 'growth'

{'CBOW': [('economy', 0.6394561529159546)], 'Skip-gram': [('economy', 0.5914967060089111)], 'GloVe': [('profits', 0.692354142665863)], 'FastText': [('profits', 0.6474031209945679)]}

-----

Query 5: 'leader' - 'male' + 'female'

{'CBOW': [('member', 0.5371192097663879)], 'Skip-gram': [('panhellenic', 0.7169508337974548)], 'GloVe': [('party', 0.7375144958496094)], 'FastText': [('leaders', 0.7392967343330383)]}

-----