In [1]: **import** pandas **as** pd from sklearn.impute import SimpleImputer # Example dataframe data = pd.read_csv('email.csv') In [2]: print(data.head()) connevey jay \ for hou 2 0 27 0 0 Email 4 0 5 22 0 5 1 51 2 10 0 0 Email 5 5 2 valued lay infrastructure military allowing ff dry Prediction 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 3 0 0 0 0 0 0 0 [5 rows x 3002 columns] In [3]: # Define the target column name target_column = 'Prediction' # Replace 'Prediction' with the actual target column name if different # Identify numerical and categorical features numerical_features = data.select_dtypes(include=['int64', 'float64']).columns.tolist() categorical_features = data.select_dtypes(include=['object']).columns.tolist() # Remove the target column from the feature lists if it's included if target_column in numerical_features: numerical_features.remove(target_column) if target_column in categorical_features: categorical_features.remove(target_column) print('Numerical Features:', numerical_features) print('Categorical Features:', categorical_features) Numerical Features: ['the', 'to', 'ect', 'and', 'for', 'of', 'a', 'you', 'hou', 'in', 'on', 'is', 'this', 'enron', 'i', 'be', 'that', 'will', 'have', 'with', 'your', 'at', 'we', 's', 'are', 'it', 'by', 'com', 'as', 'from', 'gas', 'or', 'not', 'me', 'deal', 'if', 'meter', 'hpl', 'please', 're', 'e', 'any', 'our', 'corp', 'can', 'd', 'all', 'has', 'was', 'know', 'need', 'an', 'f orwarded', 'new', 't', 'may', 'up', 'j', 'mmbtu', 'should', 'do', 'am', 'get', 'out', 'see', 'no', 'there', 'price', 'daren', 'but', 'been', 'company', 'l', 'these', 'let', 'so', 'would', 'm', 'into', 'xls', 'farmer', 'attached', 'us', 'information', 'they', 'message', 'day', 'time', 'my', 'one', 'what', 'only', 'http', 'th', 'volume', 'mail', 'contract', 'which', 'month', 'more', 'robert', 'sitara', 'about', 'texas', 'nom', 'energy', 'pec', 'questions', 'www', 'deals', 'volumes', 'pm', 'ena', 'now', 'their', 'file', 'some', 'email', 'just', 'also', 'call', 'change', 'other', 'here', 'like', 'b', 'flow', 'net', 'following', 'p', 'production', 'when', 'over', 'back', 'want', 'original', 'them', 'below', 'o', 'ticket', 'c', 'he', 'could', 'mak e', 'inc', 'report', 'march', 'contact', 'were', 'days', 'list', 'nomination', 'system', 'who', 'april', 'number', 'sale', 'don', 'its', 'first', 'thanks', 'business', 'help', 'per', 'throu gh', 'july', 'forward', 'font', 'free', 'daily', 'use', 'order', 'today', 'r', 'had', 'fw', 'set', 'plant', 'statements', 'go', 'gary', 'oil', 'line', 'sales', 'w', 'effective', 'well', 'te naska', 'take', 'june', 'x', 'within', 'nbsp', 'she', 'how', 'north', 'america', 'being', 'under', 'next', 'week', 'than', 'january', 'last', 'two', 'service', 'purchase', 'name', 'less', 'height', 'off', 'agreement', 'k', 'work', 'tap', 'group', 'year', 'based', 'transport', 'after', 'think', 'made', 'each', 'available', 'changes', 'due', 'f', 'h', 'services', 'smith', 'sen d', 'management', 'stock', 'sent', 'll', 'co', 'office', 'needs', 'cotten', 'did', 'actuals', 'u', 'money', 'before', 'looking', 'then', 'pills', 'online', 'request', 'look', 'desk', 'ami', 'his', 'same', 'george', 'chokshi', 'point', 'delivery', 'friday', 'does', 'size', 'august', 'product', 'pat', 'width', 'iv', 'noms', 'address', 'above', 'sure', 'give', 'october', 'futur e', 'find', 'market', 'n', 'mary', 'vance', 'melissa', 'said', 'internet', 'still', 'account', 'those', 'down', 'link', 'hsc', 'rate', 'people', 'pipeline', 'best', 'actual', 'very', 'end', 'home', 'houston', 'tu', 'high', 'her', 'team', 'products', 'many', 'currently', 'spot', 'receive', 'good', 'such', 'going', 'process', 'feb', 'monday', 'info', 'david', 'lloyd', 'again', 'both', 'click', 'subject', 'jackie', 'december', 'total', 'na', 'lisa', 've', 'september', 'hours', 'until', 'resources', 'because', 'aol', 'february', 'where', 'g', 'investment', 'issue', 'duke', 'since', 'pay', 'show', 'way', 'global', 'computron', 'further', 'most', 'place', 'offer', 'natural', 'activity', 'eastrans', 'graves', 'right', 'prices', 'date', 'john', 'utilitie s', 'november', 'clynes', 'jan', 'securities', 'meeting', 'susan', 'hplc', 'julie', 'able', 'received', 'align', 'term', 'id', 'revised', 'thursday', 'pg', 'fee', 'hplno', 'trading', 'addit ional', 'site', 'txu', 'data', 'wellhead', 'reply', 'taylor', 'news', 'unify', 'michael', 'provide', 'note', 'much', 'access', 'lannou', 'every', 'between', 'keep', 'tuesday', 'review', 'gr eat', 'tom', 'put', 'done', 'long', 'save', 'section', 'must', 'v', 'part', 'nd', 'million', 'check', 'trade', 'bob', 'created', 'steve', 'prior', 'copy', 'continue', 'numbers', 'via', 'wor ld', 'demand', 'hanks', 'contracts', 'phone', 'transaction', 'customer', 'possible', 'pefs', 'meyers', 'months', 'special', 'without', 'used', 'regarding', 'software', 'howard', 'support', 'buy', 'young', 'meters', 'thru', 'believe', 'gcs', 'cec', 'entered', 'control', 'dec', 'face', 'create', 'weissman', 'st', 'color', 'come', 'supply', 'brian', 'hplo', 'own', 'correct', 'cu stomers', 'web', 'allocation', 'soon', 'using', 'development', 'mark', 'low', 'power', 'problem', 'once', 'however', 'tickets', 'border', 'performance', 'manager', 'rates', 'center', 'compa nies', 'risk', 'details', 'needed', 'international', 'field', 'even', 'someone', 'doc', 'fuel', 'lee', 'paid', 'while', 'start', 'index', 'include', 'nominations', 'act', 'pricing', 'schedu led', 'gathering', 'type', 'href', 'during', 'aimee', 'anything', 'feel', 'fuels', 'getting', 'advice', 'why', 'increase', 'path', 'sell', 'works', 'issues', 'three', 'enronxgate', 'camp', 'either', 'form', 'security', 'interest', 'financial', 'family', 'xp', 'plan', 'current', 'top', 'another', 'src', 'spreadsheet', 'allen', 'wednesday', 'read', 'him', 'working', 'wynne', 'a dd', 'deliveries', 'buyback', 'allocated', 'firm', 'james', 'marketing', 'tx', 'results', 'got', 'stocks', 'calpine', 'might', 'operations', 'position', 'logistics', 'fax', 'cost', 'party', 'zero', 'pops', 'old', 'pt', 'scheduling', 'flowed', 'dollars', 'update', 'gco', 'katy', 'including', 'follow', 'yahoo', 'already', 'suite', 'error', 'past', 'page', 'stop', 'changed', 'boo k', 'program', 'few', 'better', 'operating', 'equistar', 'move', 'cotton', 'aep', 'y', 'state', 'ees', 'rita', 'provided', 'employees', 'period', 'morning', 'cd', 'hotmail', 'entex', 'swin g', 'real', 'exchange', 'tomorrow', 'lst', 'counterparty', 'parker', 'person', 'follows', 'valid', 'visit', 'little', 'professional', 'quality', 'confirm', 'something', 'megan', 'brenda', 'around', 'windows', 'im', 'storage', 'accounting', 'called', 'ranch', 'tax', 'problems', 'case', 'teco', 'fact', 'always', 'too', 'unsubscribe', 'amount', 'coastal', 'never', 'rodriguez', 'love', 'acton', 'shut', 'pipe', 'project', 'hope', 'limited', 'invoice', 'credit', 'full', 'survey', 'ray', 'carlos', 'anyone', 'wanted', 'yet', 'ic', 'scott', 'years', 'charlie', 'soft', 'notice', 'advise', 'addition', 'donald', 'lsk', 'wish', 'katherine', 'website', 'hplnl', 'schumack', 'prescription', 'cover', 'shares', 'cash', 'imbalance', 'united', 'handle', 'big', 'eve ryone', 'style', 'clear', 'producer', 'weekend', 'city', 'requested', 'stone', 'left', 'payment', 'mobil', 'shows', 'small', 'confirmed', 'technology', 'sherlyn', 'schedule', 'else', 'letter', 'box', 'bill', 'richard', 'lamphier', 'complete', 'ever', 'release', 'newsletter', 'anita', 'clem', 'having', 'herod', 'beginning', 'papayoti', 'tr y', 'mike', 'enter', 'estimates', 'location', 'cut', 'question', 'things', 'personal', 'feedback', 'cialis', 'found', 'area', 'dow', 'terms', 'central', 'necessary', 'man', 'run', 'reason', 'third', 'midcon', 'charge', 'president', 'de', 'listed', 'meds', 'thomas', 'thought', 'capital', 'added', 'ask', 'weeks', 'investing', 'commercial', 'star', 'several', 'easy', 'view', 'can not', 'extended', 'lauri', 'beaumont', 'union', 'times', 'open', 'cause', 'monthly', 'action', 'offers', 'industry', 'states', 'side', 'mailto', 'probably', 'neal', 'second', 'stephanie', 'download', 'flash', 'agree', 'mcf', 'transfer', 'doing', 'important', 'basis', 'different', 'final', 'koch', 'exxon', 'remove', 'microsoft', 'interested', 'application', 'sept', 'mg', 'wri te', 'lp', 'east', 'requirements', 'code', 'value', 'thank', 'together', 'exploration', 'mid', 'dfarmer', 'everything', 'receipt', 'thu', 'afternoon', 'late', 'enserch', 'coming', 'bank', 'response', 'tell', 'shipping', 'night', 'events', 'cynthia', 'lsp', 'close', 'legal', 'country', 'direct', 'expected', 'ces', 'corporation', 'options', 'really', 'voip', 'nominated', 'et c', 'latest', 'potential', 'priced', 'edward', 'valero', 'material', 'stack', 'victor', 'redeliveries', 'loss', 'remember', 'baumbach', 'option', 'private', 'longer', 'aware', 'included', 'drugs', 'public', 'reinhardt', 'version', 'hesse', 'discuss', 'related', 'asked', 'say', 'viagra', 'revision', 'bgcolor', 'kind', 'pro', 'completed', 'health', 'ready', 'plans', 'registere d', 'regards', 'carthage', 'zone', 'fill', 'away', 'computer', 'systems', 'industrial', 'mentioned', 'told', 'therefore', 'growth', 'sold', 'track', 'reports', 'south', 'rd', 'jim', 'cost s', 'image', 'expect', 'return', 'physical', 'el', 'browser', 'donna', 'stacey', 'begin', 'china', 'duty', 'approximately', 'showing', 'unit', 'jones', 'hard', 'verify', 'updated', 'eol', 'cs', 'orders', 'talk', 'trying', 'base', 'given', 'server', 'source', 'pathed', 'strong', 'bryan', 'directly', 'risks', 'whole', 'major', 'users', 'purchases', 'oo', 'karen', 'luong', 'lev el', 'required', 'delivered', 'portfolio', 'riley', 'ali', 'easttexas', 'poorman', 'bellamy', 'assistance', 'nothing', 'gif', 'thing', 'retail', 'didn', 'valley', 'department', 'cleburne', 'allow', 'gpgfin', 'answer', 'items', 'paste', 'avila', 'taken', 'mm', 'nguyen', 'ensure', 'reference', 'hall', 'later', 'lone', 'user', 'methanol', 'facility', 'network', 'spoke', 'thoug h', 'tabs', 'taking', 'status', 'considered', 'purchased', 'says', 'yourself', 'paliourg', 'dy', 'jeff', 'businesses', 'fred', 'transportation', 'apache', 'morris', 'nov', 'ltd', 'brand', 'federal', 'statement', 'oasis', 'reflect', 'assets', 'lamadrid', 'general', 'bridge', 'ability', 'oct', 'play', 'enrononline', 'compliance', 'spam', 'availability', 'king', 'understanding', 'chance', 'quick', 'effort', 'points', 'reliantenergy', 'fixed', 'short', 'hill', 'cheryl', 'aepin', 'key', 'understand', 'valign', 'capacity', 'game', 'took', 'bring', 'guys', 'god', 'green', 'care', 'withers', 'property', 'hub', 'johnson', 'employee', 'wants', 'albrecht', 'meaning', 'expectations', 'mx', 'moved', 'cernosek', 'matter', 'devon', 'calls', 'worldwide', 'r ecords', 'removed', 'lose', 'large', 'referenced', 'walker', 'iferc', 'enw', 'ponton', 'eileen', 'ship', 'upon', 'enerfin', 'jennifer', 'looks', 'staff', 'pc', 'target', 'waha', 'making', ecords', 'removed', 'lose', 'large', 'referenced', 'walker', 'iferc', 'enw', 'ponton', 'eileen', 'ship', 'upon', 'enerfin', 'jennifer', 'looks', 'staff', 'pc', 'target', 'waha', 'making', 'cp', 'impact', 'partner', 'immediately', 'shall', 'channel', 'takes', 'sat', 'others', 'hear', 'went', 'travel', 'listing', 'approved', 'processing', 'early', 'enough', 'sally', 'startin g', 'distribution', 'tejas', 'transactions', 'stay', 'earl', 'superty', 'doesn', 'reserves', 'includes', 'choose', 'adobe', 'publisher', 'paso', 'cornhusker', 'training', 'markets', 'conte nt', 'solution', 'shell', 'jpg', 'print', 'drive', 'pain', 'password', 'half', 'herrera', 'saturday', 'moopid', 'hotlist', 'balance', 'super', 'vacation', 'sex', 'happy', 'excess', 'existing', 'fund', 'stella', 'share', 'sign', 'wells', 'won', 'four', 'text', 'card', 'tisdale', 'fwd', 'appreciate', 'non', 'experience', 'savings', 'settlements', 'draft', 'couple', 'informe d', 'biz', 'watch', 'plus', 'sun', 'expense', 'images', 'land', 'occur', 'flowing', 'mar', 'terry', 'darren', 'cheap', 'weight', 'dynegy', 'activities', 'become', 'mr', 'format', 'attention', 'entire', 'photoshop', 'williams', 'instructions', 'neon', 'janet', 'contains', 'ago', 'friends', 'against', 'boas', 'music', 'certain', 'liz', 'svcs', 'record', 'fast', 'dave', 'held', 'mind', 'ua', 'publication', 'differ', 'comments', 'fun', 'rest', 'instant', 'agent', 'communications', 'director', 'partners', 'investors', 'expedia', 'kevin', 'assist', 'safe', 'approval', 'allocate', 'black', 'none', 'intrastate', 'document', 'eric', 'hakemack', 'expired', 'lower', 'active', 'secure', 'cc', 'five', 'determine', 'press', 'colspan', 'missing', 'jill', 'discussion', 'replication', 'secure', 'secure', 'cc', 'five', 'determine', 'press', 'doctor', 'pari 'discussion', 'relief', 'respect', 'specific', 'technologies', 'al', 'holmes', 'white', 'yesterday', 'medical', 'pinion', 'sorry', 'men', 'leave', 'pass', 'video', 'gomes', 'doctor', 'projects', 'limit', 'air', 'knle', 'pharmacy', 'confirmation', 'opportunity', 'involve', 'notify', 'gtc', 'class', 'ken', 'started', 'outage', 'confidential', 'room', 'blue', 'estimated', 'off icer', 'reach', 'messages', 'database', 'words', 'prc', 'tracked', 'transition', 'light', 'national', 'hot', 'offering', 'gulf', 'provides', 'iit', 'demokritos', 'mckay', 'average', 'wid e', 'heard', 'files', 'dan', 'billed', 'mccoy', 'rc', 'exactly', 'middle', 'select', 'bruce', 'louisiana', 'receiving', 'california', 'event', 'roll', 'mops', 'william', 'appear', 'perfec t', 'html', 'features', 'join', 'greater', 'sunday', 'pick', 'featured', 'cdnow', 'prize', 'reveffo', 'olsen', 'expects', 'estimate', 'near', 'common', 'package', 'title', 'whether', 'boug ht', 'evergreen', 'difference', 'elizabeth', 'history', 'monitor', 'advised', 'result', 'sources', 'school', 'unaccounted', 'paragraph', 'turn', 'kimberly', 'increased', 'communication', 'members', 'concerns', 'uncertainties', 'associated', 'reduce', 'committed', 'wi', 'asap', 'goes', 'trader', 'waiting', 'canada', 'worth', 'representative', 'claim', 'ceo', 'london', 'disc ussions', 'php', 'brazos', 'trevino', 'calling', 'involved', 'la', 'gift', 'southern', 'groups', 'hour', 'tufco', 'previously', 'voice', 'normally', 'resolve', 'efforts', 'nor', 'recent', 'purchasing', 'county', 'ok', 'express', 'generic', 'according', 'respond', 'situation', 'hold', 'lot', 'interconnect', 'word', 'came', 'west', 'role', 'opportunities', 'corporate', 'remai n', 'similar', 'readers', 'suggestions', 'subscribers', 'projections', 'lead', 'learn', 'resolved', 'agreed', 'sec', 'head', 'enjoy', 'img', 'rnd', 'responsible', 'outstanding', 'member', 'panenergy', 'american', 'cass', 'register', 'promotions', 'parties', 'winfree', 'selling', 'usage', 'appropriate', 'assignment', 'media', 'believes', 'require', 'submit', 'model', 'spinna ker', 'copano', 'facilities', 'opinion', 'factors', 'identified', 'beverly', 'ews', 'gdp', 'deliver', 'job', 'profile', 'across', 'neuweiler', 'suggest', 'girls', 'manage', 'usa', 'local', 'bad', 'greg', 'vs', 'fees', 'digital', 'cf', 'strangers', 'registration', 'delta', 'rolex', 'goliad', 'hesco', 'success', 'primary', 'quarter', 'course', 'chairman', 'petroleum', 'notes', 'medications', 'ei', 'instead', 'fine', 'lake', 'pre', 'force', 'seek', 'recipient', 'gain', 'placed', 'age', 'least', 'body', 'asking', 'discussed', 'hanson', 'emails', 'nominate', 'ext', 'known', 'ones', 'ed', 'assigned', 'htmlimg', 'means', 'present', 'various', 'invoices', 'gd', 'agency', 'along', 'located', 'reflects', 'solutions', 'ex', 'house', 'cds', 'br', 'owner', 'apr', 'sullivan', 'basin', 'linda', 'worked', 'car', 'seen', 'properties', 'booked', 'higher', 'store', 'est', 'revenue', 'wait', 'women', 'far', 'met', 'wholesale', 'range', 'kcs', 'reco rded', 'brown', 'lots', 'match', 'input', 'grant', 'providing', 'huge', 'investor', 'kelly', 'apply', 'paths', 'handling', 'pipes', 'advantage', 'analysis', 'focus', 'draw', 'red', 'origin ation', 'connection', 'planning', 'wilson', 'golf', 'summary', 'item', 'bankruptcy', 'expenses', 'pgev', 'encina', 'beaty', 'memo', 'initial', 'thousand', 'mills', 'penis', 'friend', 'conv ersation', 'multiple', 'martin', 'names', 'bit', 'dth', 'talked', 'behalf', 'preliminary', 'button', 'herein', 'gisb', 'coupon', 'sa', 'oi', 'appears', 'door', 'texaco', 'csikos', 'arrange ments', 'cpr', 'expires', 'popular', 'sending', 'research', 'conditions', 'gb', 'board', 'ca', 'applications', 'tried', 'paying', 'acquisition', 'reporting', 'normal', 'maintenance', 'resu me', 'announced', 'attachment', 'buyer', 'objectives', 'prod', 'represent', 'sandi', 'hplnol', 'government', 'committee', 'running', 'tetco', 'discount', 'jo', 'holding', 'earlier', 'posit ions', 'happen', 'mailing', 'decided', 'recently', 'chris', 'xanax', 'valium', 'broadband', 'individual', 'station', 'td', 'financing', 'somehow', 'pena', 'critical', 'attend', 'kristen', ions', 'happen', 'mailing', 'decided', 'recently', 'chris', 'xanax', 'valium', 'broadband', 'individual', 'station', 'td', 'financing', 'somehow', 'pena', 'critical', 'attend', 'kristen', 'inform', 'highly', 'hl', 'phillips', 'minutes', 'titles', 'affiliate', 'wife', 'lonestar', 'charlotte', 'quickly', 'paper', 'test', 'comes', 'mobile', 'internal', 'privacy', 'ideas', 'live', 'gotten', 'floor', 'benefit', 'percent', 'ms', 'dr', 'ebs', 'msn', 'gave', 'dallas', 'enterprise', 'rx', 'spring', 'ftar', 'ooking', 'hawkins', 'exclusive', 'selected', 'baxter', 'actually', 'single', 'shop', 'nominates', 'guarantee', 'minute', 'correctly', 'unique', 'bid', 'building', 'stated', 'accept', 'assumptions', 'centana', 'senior', 'pill', 'kinsey', 'sap', 'immediate', 'goals', 'category', 'mitchell', 'acceptance', 'termination', 'sweeney', 'facts', 'amazon', 'arrangement', 'josey', 'funds', 'among', 'accuracy, 'mean', 'rather', 'kim', 'egmnom', 'indicate', 'updates', 'extra', 'adjustment', 'accounts', 'gold', 'purposes', 'remaining', 'talking', 'entry', 'road', 'load', 'simply', 'europe', 'lindley', 'understood', 'logos', 'hi', 'speed', 'profit', 'notified', 'jackson', 'z', 'vols', 'serve', 'additionally', 'shipped', 'connor', 'fontfont', 'q', 'kept', 'dollar', 'jr', 'almost', 'fri', 'paul', 'documents', 'analyst', 'crude', 'cap', 'shopping', 'aug', 'clearance', 'schneider', 'ftworth', 'father', 'anticipated', 'resellers', 'conterprises', 'counterprises', 'epgt', 'buying', 'san', 'invest', 'cartwheel', 'brandywine', 'wrong', 'mtbe', 'split', 'submitted', 'hull', 'gra', 'children', 'leader', 'TRUE', 'baseload', 'mb', 'letters', 'billion', 'rights', 'mtr', 'heidi', 'clean', 'historical'. 'asset'. 'foreign'. 'gra', 'entity'. 'developed'. 'maybe'. 'ieffrey'. 'transmission'. 'outside'. 'lost'. 'membership', 'invitation', 'locean', 'legislation', 'hernandez', 'pep', 'historical', 'asset', 'foreign', 'gr', 'entity', 'developed', 'maybe', 'jeffrey', 'transmission', 'outside', 'lost', 'membership', 'invitation', 'ocean', 'legislation', 'hernandez', 'pep', 'payments', 'wallis', 'rev', 'kenneth', 'seaman', 'annual', 'guess', 'bammel', 'lines', 'guadalupe', 'zivley', 'exception', 'example', 'pathing', 'revisions', 'pipelines', 'equity', 'budge t', 'wed', 'dealers', 'window', 'juno', 'claims', 'bottom', 'standard', 'alternative', 'merchant', 'braband', 'topica', 'telephone', 'reliant', 'speculative', 'yes', 'en', 'morgan', 'cabl e', 'edmondson', 'participate', 'usb', 'throughout', 'checked', 'myself', 'contents', 'fat', 'investments', 'six', 'build', 'giving', 'calendar', 'inherent', 'edition', 'darial', 'hr', 'tri 'pull', 'moving', 'concern', 'proposed', 'rm', 'deer', 'enquiries', 'alt', 'tammy', 'front', 'reduction', 'evening', 'concerning', 'gets', 'effect', 'isn', 'haven', 'cowboy', 'sea', 'dv 'launch', 'minimum', 'changing', 'built', 'avoid', 'chief', 'stephen', 'chad', 'manual', 'finally', 'strategy', 'executive', 'thousands', 'conflict', 'resulting', 'policy', 'commissio n', 'stand', 'positive', 'quantity', 'programs', 'airmail', 'texoma', 'prepared', 'austin', 'matt', 'intent', 'uae', 'citibank', 'jaquet', 'hol', 'harris', 'min', 'hplr', 'advance', 'weathe 'terminated', 'whom', 'sheet', 'venturatos', 'cellpadding', 'hotel', 'leading', 'guaranteed', 'idea', 'announce', 'pleased', 'award', 'operational', 'prepare', 'schedulers', 'child', 's um', 'quote', 'adjusted', 'warning', 'issued', 'ga', 'cross', 'detail', 'pertaining', 'tess', 'owe', 'crow', 'availabilities', 'griffin', 'christy', 'crosstex', 'eel', 'itoy', 'heart', 'lic ensed', 'overnight', 'cal', 'otherwise', 'luck', 'stretch', 'generation', 'broker', 'construed', 'except', 'traders', 'carry', 'column', 'approx', 'main', 'alert', 'charges', 'step', 'reven ues', 'games', 'gottlob', 'looked', 'individuals', 'beck', 'stuff', 'welcome', 'port', 'glover', 'description', 'daniel', 'quantities', 'park', 'managing', 'town', 'seller', 'summer', 'tin a', 'dates', 'eff', 'dudley', 'ferc', 'robin', 'charles', 'customerservice', 'zonedubai', 'emirates', 'aeor', 'clickathome', 'materia', 'island', 'vaughn', 'sexual', 'eiben', 'forms', 'dele te', 'realize', 'tailgate', 'behind', 'villarreal', 'lon', 'benoit', 'simple', 'tech', 'ahead', 'double', 'ordering', 'se', 'miss', 'law', 'eb', 'post', 'outlook', 'equipment', 'leslie', 'r eeves', 'org', 'tools', 'cold', 'adjustments', 'contained', 'saw', 'edit', 'deciding', 'finance', 'patti', 'listbot', 'river', 'kathryn', 'holiday', 'successful', 'unable', 'advisor', 'poo l', 'bryce', 'outages', 'adjust', 'screen', 'otc', 'brent', 'helps', 'auto', 'foot', 'region', 'links', 'contain', 'knowledge', 'yvette', 'dial', 'pressure', 'detailed', 'indicated', 'charg ed', 'sites', 'makes', 'female', 'mcmills', 'cook', 'mazowita', 'meredith', 'allocations', 'meetings', 'particular', 'environment', 'drug', 'search', 'mailings', 'designed', 'rock', 'measur , 'corrected', 'kids', 'benefits', 'tv', 'seems', 'husband', 'fix', 'grow', 'decision', 'wireless', 'mo', 'conference', 'interview', 'levels', 'copies', 'cindy', 'urgent', 'reg ular', 'payroll', 'shown', 'consumers', 'reliable', 'tr', 'indicating', 'coast', 'greif', 'severson', 'tri', 'vicodin', 'liquids', 'significant', 'intend', 'usd', 'pager', 'avails', 'spence r', 'ce', 'charset', 'verdana', 'fully', 'flynn', 'da', 'personnel', 'multi', 'closed', 'vice', 'administration', 'gmt', 'midstream', 'eye', 'speckels', 'studio', 'cilco', 'likely', 'manage rs', 'structure', 'sit', 'parent', 'preparation', 'mix', 'mmbtus', 'timing', 'happening', 'lottery', 'killing', 'acquire', 'mack', 'pcx', 'fares', 'internationa', 'notification', 'swift', 'identify', 'areas', 'separate', 'unless', 'producers', 'allows', 'pretty', 'waste', 'joanie', 'drop', 'taxes', 'premium', 'teams', 'choice', 'largest', 'addressed', 'dolphin', 'ngo', 'sel f', 'davis', 'htm', 'ad', 'graphics', 'hit', 'competitive', 'thus', 'incorrect', 'ti', 'acts', 'previous', 'edu', 'proven', 'electric', 'pictures', 'charlene', 'benedict', 'chevron', 'treat ment', 'lesson', 'player', 'sds', 'wc', 'intraday', 'assurance', 'sdsnom', 'rebecca', 'quit', 'netco', 'intra', 'whatever', 'lyondell', 'reviewed', 'solicitation', 'filings', 'log', 'noon', 'locations', 'joe', 'completely', 'rivers', 'language', 'street', 'automatically', 'ft', 'powerful', 'specials', 'alone', 'fyi', 'properly', 'proper', 'explode', 'decrease', 'medication', 'desks', 'impacted', 'anywhere', 'completion', 'banking', 'consider', 'certificate', 'zeroed', 'websites', 'tonight', 'diligence', 'education', 'club', 'vegas', 'affordable', 's ports', 'predictions', 'billing', 'diamond', 'posted', 'prayer', 'actions', 'nomad', 'resuits', 'jason', 'purpose', 'deposit', 'entertainment', 'materially', 'blank', 'resolution', 'anderso n', 'nat', 'rom', 'soma', 'organization', 'aquila', 'solid', 'affected', 'transco', 'spend', 'responsibilities', 'assume', 'header', 'accountant', 'functionality', 'meant', 'killed', 'analy sts', 'rick', 'rolled', 'noted', 'discovered', 'offices', 'torch', 'often', 'york', 'joint', 'briley', 'competition', 'guide', 'intercompany', 'son', 'settlement', 'presently', 'cart', 'ti m', 'entries', 'russ', 'valadez', 'rules', 'molly', 'apple', 'atleast', 'scheduler', 'pi', 'hector', 'dell', 'opm', 'hottlist', 'yap', 'gone', 'heal', 'llc', 'setting', 'reached', 'proposa l', 'hundred', 'trust', 'official', 'table', 'mcgee', 'written', 'operation', 'cellspacing', 'laptop', 'feature', 'ram', 'victoria', 'larry', 'units', 'requests', 'continued', 'external', 'pack', 'couldn', 'lateral', 'strictly', 'resource', 'although', 'sr', 'commodity', 'pulled', 'protocol', 'bed', 'generated', 'redmond', 'girl', 'apparently', 'tool', 'reviews', 'released', 'movies', 'inside', 'shareholder', 'rr', 'compensation', 'beliefs', 'foresee', 'lease', 'rule', 'marta', 'chemical', 'hillary', 'hp', 'tongue', 'advises', 'master', 'eight', 'was n', 'itself', 'documentation', 'xl', 'humble', 'elsa', 'pics', 'hughes', 'brokered', 'distribute', 'consultation', 'sheri', 'lists', 'cannon', 'treated', 'factor', 'putting', 'verified', 'r eleases', 'enhanced', 'controls', 'craig', 'worksheet', 'conversion', 'max', 'hrs', 'helpful', 'hand', 'producing', 'dl', 'developing', 'design', 'woman', 'understands', 'standards', 'promo tion', 'sarco', 'hospital', 'ffffff', 'respective', 'richmond', 'conoco', 'driver', 'easily', 'sean', 'den', 'gateway', 'holdings', 'brad', 'college', 'gains', 'adult', 'dated', 'em', 'mclo ughlin', 'anticipates', 'henderson', 'julia', 'negotiations', 'sofftwaares', 'garrick', 'comstock', 'trochta', 'imceanotes', 'ecom', 'larger', 'nommensen', 'coordinate', 'partnership', 'otc bb', 'announces', 'louis', 'dealer', 'reliance', 'season', 'agua', 'dulce', 'offshore', 'gathered', 'forever', 'function', 'happened', 'sample', 'easier', 'aim', 'pa', 'expensive', 'think s', 'maximum', 'war', 'mining', 'drilling', 'owned', 'todd', 'advanced', 'provider', 'pending', 'providers', 'silver', 'cherry', 'hundreds', 'thoughts', 'addresses', 'beach', 'baby', 'requi res', 'caused', 'variance', 'extension', 'carbide', 'anytime', 'adding', 'triple', 'dawn', 'martinez', 'entering', 'login', 'bretz', 'ls', 'writeoff', 'locker', 'wiil', 'block', 'blood', 'r omeo', 'responsibility', 'brennan', 'btu', 'venture', 'connected', 'nascar', 'opinions', 'executed', 'cell', 'flag', 'doctors', 'invoiced', 'marlin', 'coffey', 'nice', 'amazing', 'ii', 'det ermined', 'handled', 'keeping', 'touch', 'upgrade', 'shipment', 'brought', 'forwarding', 'confidence', 'hesitate', 'seem', 'electronic', 'appreciated', 'deadline', 'franklin', 'heather', 'r easons', 'passed', 'safety', 'procedures', 'payback', 'networks', 'utility', 'count', 'africa', 'exact', 'creating', 'loading', 'processed', 'court', 'tier', 'sender', 'att', 'mailbox', 'gl ad', 'buddy', 'profiles', 'portion', 'protection', 'compressor', 'okay', 'oba', 'finding', 'heads', 'bar', 'turned', 'remote', 'illustrator', 'oem', 'noticed', 'mails', 'darron', 'nick', 'u rbanek', 'jerry', 'barrett', 'ehronline', 'und', 'abdv', 'egm', 'couid', 'technoiogies', 'owns', 'improved', 'eat', 'moment', 'owners', 'develop', 'installed', 'videos', 'frank', 'hearing', 'inches', 'busy', 'ref', 'valuable', 'et', 'un', 'url', 'shawna', 'iso', 'capture', 'extremely', 'ya', 'causing', 'consent', 'anyway', 'round', 'discrepancies', 'cheapest', 'confidentialit y', 'disclosure', 'prohibited', 'vol', 'correction', 'communicate', 'processes', 'spain', 'shareholders', 'supported', 'smoking', 'mine', 'biggest', 'erections', 'platform', 'miles', 'excit ing', 'association', 'die', 'restricted', 'ma', 'income', 'goal', 'bane', 'collection', 'nathan', 'wind', 'piece', 'familiar', 'gore', 'experiencing', 'pico', 'mai', 'dewpoint', 'tessie', 'hair', 'bussell', 'diane', 'delivering', 'originally', 'accurate', 'began', 'seven', 'tracking', 'randall', 'gay', 'emerging', 'prescriptions', 'story', 'arial', 'florida', 'space', 'owner ship', 'european', 'sutton', 'concerned', 'male', 'spent', 'agreements', 'industries', 'picture', 'filled', 'continues', 'death', 'choate', 'majeure', 'device', 'hence', 'ten', 'campaign', 'massive', 'eyes', 'requesting', 'lives', 'reminder', 'eliminate', 'copied', 'consemiu', 'died', 'sound', 'offered', 'expressed', 'anti', 'duplicate', 'steps', 'books', 'improve', 'implemen tation', 'gives', 'ac', 'peggy', 'proprietary', 'ways', 'advertisement', 'published', 'earnings', 'mortgage', 'consumer', 'ct', 'tape', 'fl', 'cia', 'organizational', 'agenda', 'rental', 'c arriere', 'moshou', 'church', 'trouble', 'medium', 'aggressive', 'smart', 'zajac', 'ail', 'participants', 'gap', 'earthlink', 'wire', 'trades', 'messaging', 'ut', 'wil', 'richardson', 'blv d', 'glo', 'seneca', 'pubiisher', 'imited', 'isc', 'contacts', 'sleep', 'kyle', 'cooperation', 'possibly', 'leaving', 'motor', 'hopefully', 'tie', 'speak', 'mi', 'suggested', 'canadian', 'u ses', 'connect', 'pvr', 'rich', 'places', 'auction', 'po', 'spacer', 'client', 'recommended', 'default', 'living', 'regardless', 'human', 'bringing', 'focused', 'store s', 'variety', 'netherlands', 'leaders', 'bowen', 'salary', 'signed', 'penny', 'loan', 'desktop', 'chase', 'pleasure', 'compare', 'session', 'overall', 'stranger', 'length', 'planned', 's p', 'darrel', 'raise', 'palestinian', 'expiration', 'serial', 'premiere', 'suzanne', 'reduced', 'players', 'applicable', 'impotence', 'buckley', 'wayne', 'hansen', 'indicative', 'sabrae', 'dating', 'winners', 'marshall', 'highest', 'ea', 'presentation', 'allowed', 'square', 'danny', 'gepl', 'hydrocarbon', 'alpine', 'christmas', 'muscle', 'souza', 'relating', 'begins', 'ecf', 'forth', 'answers', 'audit', 'approve', 'lunch', 'types', 'starts', 'difficult', 'le', 'lasts', 'series', 'till', 'edge', 'growing', 'covered', 'shipper', 'sometime', 'republic', 'filter', 'sooner', 'increasing', 'nelson', 'percentage', 'returned', 'pop', 'interface', 'kin', 'experienced', 'prime', 'merger', 'obtain', 'ryan', 'servers', 'attachments', 'achieve', 'effects', 'g ov', 'examples', 'procedure', 'explore', 'caribbean', 'rally', 'amounts', 'comfort', 'attempt', 'greatly', 'amelia', 'engel', 'delay', 'fare', 'der', 'cove', 'filing', 'fletcher', 'leth', 'undervalued', 'cents', 'esther', 'hlavaty', 'reid', 'lls', 'troy', 'palmer', 'metals', 'las', 'carter', 'luis', 'migration', 'brief', 'hess', 'therein', 'ur', 'pond', 'joanne', 'communit y', 'tglo', 'eogi', 'ml', 'wysak', 'felipe', 'errors', 'affect', 'convenient', 'minimal', 'boost', 'incremental', 'decide', 'reserve', 'superior', 'kerr', 'willing', 'quite', 'wild', 'unlim ited', 'sans', 'mother', 'computers', 'unfortunately', 'ordered', 'satisfaction', 'priority', 'traded', 'testing', 'portal', 'ward', 'lets', 'aren', 'knows', 'refer', 'shot', 'fda', 'tue', 'saying', 'cancel', 'forecast', 'cousino', 'bass', 'permanent', 'phones', 'technical', 'whose', 'objective', 'cards', 'distributed', 'learning', 'fire', 'drill', 'towards', 'forget', 'explo sion', 'gloria', 'formula', 'redelivery', 'audio', 'visual', 'encoding', 'approach', 'doubt', 'staffing', 'excite', 'corel', 'tm', 'enronavailso', 'contacting', 'alland', 'heavy', 'economi c', 'nigeria', 'milwaukee', 'phillip', 'curve', 'returns', 'padre', 'kathy', 'buttons', 'sir', 'vary', 'sounds', 'disclose', 'authority', 'flw', 'straight', 'worldnet', 'beemer', 'ooo', 'de fs', 'thorough', 'officers', 'flight', 'prefer', 'awesome', 'macintosh', 'feet', 'constitutes', 'formosa', 'porn', 'armstrong', 'driscoll', 'watches', 'newsietter', 'twenty', 'tommy', 'fiel ds', 'method', 'setup', 'allocating', 'initially', 'missed', 'clarification', 'especially', 'dorcheus', 'del', 'millions', 'insurance', 'pooling', 'trial', 'tennessee', 'ellis', 'direction', 'setup', 'allocating', 'initially', 'missed', 'clarification', 'especially', 'dorcheus', 'del', 'millions', 'insurance', 'pooling', 'trial', 'tennessee', 'ellis', 'direction', 'setup', 'allocating', 'missed', 'clarification', 'especially', 'dorcheus', 'del', 'millions', 'insurance', 'pooling', 'trial', 'tennessee', 'ellis', 'direction', 'setup', 'missed', 'clarification', 'especially', 'dorcheus', 'del', 'millions', 'insurance', 'pooling', 'trial', 'tennessee', 'ellis', 'direction', 'setup', 'missed', 'clarification', 'especially', 'dorcheus', 'del', 'millions', 'insurance', 'pooling', 'trial', 'tennessee', 'ellis', 'direction', 'especially', 'dorcheus', 'del', 'missed', 'trial', 'tria n', 'bold', 'catch', 'performing', 'accepted', 'matters', 'batch', 'continuing', 'winning', 'symbol', 'offsystem', 'decisions', 'produced', 'ended', 'greatest', 'degree', 'solmonson', 'imba lances', 'fall', 'fear', 'hate', 'fight', 'reallocated', 'debt', 'reform', 'australia', 'plain', 'prompt', 'remains', 'ifhsc', 'enhancements', 'connevey', 'jay', 'valued', 'lay', 'infrastru cture', 'military', 'allowing', 'ff', 'dry'] Categorical Features: ['Email No.'] Handle Missing Values missing_values = data.isnull().sum() print("Missing values in each column:\n", missing_values) Missing values in each column: Email No. ect military allowing Prediction Length: 3002, dtype: int64 **Encode Categorical Variables** In [5]: # One-Hot Encoding for nominal features (if any) data = pd.get_dummies(data, columns=categorical_features, drop_first=True) print('Categorical features encoded.') print(data.head()) Categorical features encoded. the to ect and for of a you hou in ... Email No._Email 990 0 2 0 0 0 ... 2 102 27 18 ... 6 5 51 2 10 1 ... 2 57 Email No._Email 991 Email No._Email 992 Email No._Email 993 \ 0 0 0 1 2 0 0 0 0 3 Email No._Email 994 Email No._Email 995 Email No._Email 996 \ 1 0 0 Email No._Email 997 Email No._Email 998 Email No._Email 999 0 2 0 [5 rows x 8172 columns] Scale Numerical Features In [6]: **from** sklearn.preprocessing **import** StandardScaler # Scaling Numerical Features scaler = StandardScaler() data[numerical_features] = scaler.fit_transform(data[numerical_features]) print('Numerical features scaled.') print(data.head()) Numerical features scaled. the to ect and for of 0 -0.565449 -0.649083 -0.293895 -0.508752 -0.667663 -0.421725 -0.611169 1 0.115757 0.714508 1.337337 0.483741 0.614369 -0.100659 0.530831 2 -0.565449 -0.649083 -0.293895 -0.508752 -0.667663 -0.421725 -0.542649 3 -0.565449 -0.124625 1.195490 -0.508752 0.400697 -0.261192 -0.051589 4 0.030606 -0.019733 0.840875 -0.343336 0.400697 -0.100659 0.016931 in ... Email No._Email 990 ∖ hou you 0 -0.571751 -0.290556 -0.549800 ... 1 -0.339949 3.584743 0.383809 ... 2 -0.571751 -0.290556 -0.342331 ... 3 -0.108147 1.144740 -0.497933 ... 4 -0.571751 1.001210 -0.394198 ... Email No._Email 991 Email No._Email 992 Email No._Email 993 \ Email No._Email 994 Email No._Email 995 Email No._Email 996 \ 0 0 0 1 2 0 0 0 3 0 0 0 4 Email No._Email 997 Email No._Email 998 Email No._Email 999 0 0 0 1 0 0 [5 rows x 8172 columns] Split the Dataset In [7]: from sklearn.model_selection import train_test_split # Splitting data into features and target X = data.drop(target_column, axis=1) y = data[target_column] # Split the dataset into training and testing sets X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42) print('Dataset split into training and testing sets.') print('X_train shape:', X_train.shape) print('X_test shape:', X_test.shape) Dataset split into training and testing sets. X_train shape: (4137, 8171) X_test shape: (1035, 8171) **Model Selection and Training** In [8]; from sklearn.linear_model import LogisticRegression # Model Selection: Example with Logistic Regression log_reg = LogisticRegression() log_reg.fit(X_train, y_train) print('Logistic Regression model trained.') Logistic Regression model trained. C:\ProgramData\anaconda3\Lib\site-packages\sklearn\linear_model_logistic.py:460: ConvergenceWarning: lbfgs failed to converge (status=1): STOP: TOTAL NO. of ITERATIONS REACHED LIMIT. Increase the number of iterations (max_iter) or scale the data as shown in: https://scikit-learn.org/stable/modules/preprocessing.html Please also refer to the documentation for alternative solver options: https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression n_iter_i = _check_optimize_result(In [9]: **from** sklearn.tree **import** DecisionTreeClassifier # Model Selection: Decision Tree decision_tree = DecisionTreeClassifier() decision_tree.fit(X_train, y_train) print('Decision Tree model trained.') Decision Tree model trained. In [10]: **from** sklearn.svm **import** SVC # Model Selection: Support Vector Machine svm = SVC()svm.fit(X_train, y_train) print('Support Vector Machine model trained.') Support Vector Machine model trained. Model Evaluation In [11]: from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score # Function to evaluate model performance def evaluate_model(model, X_test, y_test): y_pred = model.predict(X_test) accuracy = accuracy_score(y_test, y_pred) precision = precision_score(y_test, y_pred) recall = recall_score(y_test, y_pred) f1 = f1_score(y_test, y_pred) return accuracy, precision, recall, f1 # Evaluate Logistic Regression log_reg_metrics = evaluate_model(log_reg, X_test, y_test) print(f'Logistic Regression - Accuracy: {log_reg_metrics[0]}, Precision: {log_reg_metrics[1]}, Recall: {log_reg_metrics[2]}, F1-score: {log_reg_metrics[3]}') # Evaluate Decision Tree decision_tree_metrics = evaluate_model(decision_tree, X_test, y_test) print(f'Decision Tree - Accuracy: {decision_tree_metrics[0]}, Precision: {decision_tree_metrics[1]}, Recall: {decision_tree_metrics[2]}, F1-score: {decision_tree_metrics[3]}') # Evaluate SVM svm_metrics = evaluate_model(svm, X_test, y_test) print(f'Support Vector Machine - Accuracy: {svm_metrics[0]}, Precision: {svm_metrics[1]}, Recall: {svm_metrics[2]}, F1-score: {svm_metrics[3]}') Logistic Regression - Accuracy: 0.9719806763285024, Precision: 0.9348534201954397, Recall: 0.9695945945945946, F1-score: 0.9519071310116085 Decision Tree - Accuracy: 0.9227053140096618, Precision: 0.8648648648649, Recall: 0.8648648648649, F1-score: 0.8648648648648649 Support Vector Machine - Accuracy: 0.9468599033816425, Precision: 0.9958847736625515, Recall: 0.8175675675675675, F1-score: 0.8979591836734693 **Cross-Validation** In [12]: from sklearn.model_selection import cross_val_score # Function to perform cross-validation and return the mean accuracy def cross_validate_model(model, X, y): cross_val_scores = cross_val_score(model, X, y, cv=5, scoring='accuracy') return cross_val_scores.mean() # Cross-validate Logistic Regression log_reg_cv_score = cross_validate_model(log_reg, X, y) print(f'Logistic Regression Cross-validated accuracy: {log_reg_cv_score}') # Cross-validate Decision Tree decision_tree_cv_score = cross_validate_model(decision_tree, X, y) print(f'Decision Tree Cross-validated accuracy: {decision_tree_cv_score}') # Cross-validate SVM svm_cv_score = cross_validate_model(svm, X, y) print(f'Support Vector Machine Cross-validated accuracy: {svm_cv_score}') C:\ProgramData\anaconda3\Lib\site-packages\sklearn\linear_model_logistic.py:460: ConvergenceWarning: lbfgs failed to converge (status=1): STOP: TOTAL NO. of ITERATIONS REACHED LIMIT. Increase the number of iterations (max_iter) or scale the data as shown in: https://scikit-learn.org/stable/modules/preprocessing.html Please also refer to the documentation for alternative solver options: https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression n_iter_i = _check_optimize_result(C:\ProgramData\anaconda3\Lib\site-packages\sklearn\linear_model_logistic.py:460: ConvergenceWarning: lbfgs failed to converge (status=1): Increase the number of iterations (max_iter) or scale the data as shown in: https://scikit-learn.org/stable/modules/preprocessing.html Please also refer to the documentation for alternative solver options: https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression n_iter_i = _check_optimize_result(C:\ProgramData\anaconda3\Lib\site-packages\sklearn\linear_model_logistic.py:460: ConvergenceWarning: lbfgs failed to converge (status=1): STOP: TOTAL NO. of ITERATIONS REACHED LIMIT. Increase the number of iterations (max_iter) or scale the data as shown in: https://scikit-learn.org/stable/modules/preprocessing.html Please also refer to the documentation for alternative solver options: https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression n_iter_i = _check_optimize_result(C:\ProgramData\anaconda3\Lib\site-packages\sklearn\linear_model_logistic.py:460: ConvergenceWarning: lbfgs failed to converge (status=1): STOP: TOTAL NO. of ITERATIONS REACHED LIMIT. Increase the number of iterations (max_iter) or scale the data as shown in: https://scikit-learn.org/stable/modules/preprocessing.html Please also refer to the documentation for alternative solver options: https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression n_iter_i = _check_optimize_result(C:\ProgramData\anaconda3\Lib\site-packages\sklearn\linear_model_logistic.py:460: ConvergenceWarning: lbfgs failed to converge (status=1): STOP: TOTAL NO. of ITERATIONS REACHED LIMIT. Increase the number of iterations (max_iter) or scale the data as shown in: https://scikit-learn.org/stable/modules/preprocessing.html Please also refer to the documentation for alternative solver options: https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression n_iter_i = _check_optimize_result(Logistic Regression Cross-validated accuracy: 0.9592023846232912 Decision Tree Cross-validated accuracy: 0.9170515515936423 Support Vector Machine Cross-validated accuracy: 0.9211100832562442 Ensemble Methods and Hyperparameter Tuning from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassifier from sklearn.model_selection import GridSearchCV, RandomizedSearchCV # Ensemble Methods: Random Forests and Gradient Boosting random_forest = RandomForestClassifier() gradient_boosting = GradientBoostingClassifier() random_forest.fit(X_train, y_train) gradient_boosting.fit(X_train, y_train) print('Ensemble methods trained.') Ensemble methods trained. In [14]: # Hyperparameter Tuning: Grid Search Example param_grid = {'C': [0.1, 1, 10], 'solver': ['liblinear']} grid_search = GridSearchCV(LogisticRegression(), param_grid, cv=5) grid_search.fit(X_train, y_train) print('Best parameters found by Grid Search:', grid_search.best_params_) Best parameters found by Grid Search: {'C': 0.1, 'solver': 'liblinear'} In [15]: # Random Search Example param_dist = {'n_estimators': [10, 50, 100], 'max_depth': [None, 10, 20, 30]} random_search = RandomizedSearchCV(RandomForestClassifier(), param_distributions=param_dist, n_iter=10, cv=5) random_search.fit(X_train, y_train) print('Best parameters found by Random Search:', random_search.best_params_) Best parameters found by Random Search: {'n_estimators': 100, 'max_depth': None}