

Name: Bharath Palanisamy Mettukadai

Reg No: 23MID0014

```
In [9]: def gender_features(word):  
        return {'last_letter':word[-1]}  
  
gender_features('Obama')
```

Out[9]: {'last_letter': 'a'}

```
In [10]: from nltk.corpus import names  
names.words()  
print(len(names.words()))  
  
labeled_names = [(name, 'male') for name in names.words('male.txt')] + [(name, 'female') for name in names.words('female.txt')]  
  
import random  
random.shuffle(labeled_names)  
  
featuresets = [(gender_features(n), gender) for (n,gender) in labeled_names]  
#Try Printing featuresets  
  
train_set, test_set = featuresets[5000:], featuresets[:2000]  
  
import nltk  
classifier = nltk.NaiveBayesClassifier.train(train_set)  
  
classifier.classify(gender_features('David'))  
#classifier.classify(gender_features('Yann Lecun'))  
  
#print(nltk.classify.accuracy(classifier, test_set))
```

7944

Out[10]: 'male'

```
In [11]: labeled_names
```

```
Out[11]: [('Joab', 'male'),
          ('Marthe', 'female'),
          ('Brynn', 'female'),
          ('Mohammed', 'male'),
          ('Sidonia', 'female'),
          ('Wandis', 'female'),
          ('Shira', 'female'),
          ('Misty', 'female'),
          ('Ryann', 'female'),
          ('Averil', 'female'),
          ('Walt', 'male'),
          ('Eba', 'female'),
          ('Kylan', 'female'),
          ('Sammy', 'female'),
          ('Evette', 'female'),
          ('Kalila', 'female'),
          ('Theadora', 'female'),
          ('Athene', 'female'),
          ('Christin', 'female'),
          ('Coraline', 'female'),
          ('Tracy', 'male'),
          ('Devon', 'male'),
          ('Katee', 'female'),
          ('Aldo', 'male'),
          ('Lyndy', 'female'),
          ('Theresita', 'female'),
          ('Leif', 'male'),
          ('Codie', 'female'),
          ('Anette', 'female'),
          ('Ulrica', 'female'),
          ('Cacilie', 'female'),
          ('Carrie', 'female'),
          ('Nat', 'female'),
          ('Grissel', 'female'),
          ('Berte', 'female'),
          ('Lottie', 'female'),
          ('Alleen', 'female'),
          ('Jeni', 'female'),
          ('Darya', 'female'),
          ('Hanni', 'female'),
          ('Giraud', 'male'),
          ('Olimpia', 'female'),
          ('Birdie', 'female'),
          ('Shepard', 'male'),
          ('Georgy', 'male'),
          ('Mathilde', 'female'),
          ('Daphene', 'female'),
          ('Ahmet', 'male'),
          ('Benton', 'male'),
          ('Elaina', 'female'),
          ('Venkat', 'male'),
          ('Christal', 'female'),
          ('Sinead', 'female'),
          ('Simmonds', 'male'),
          ('Fleming', 'male'),
          ('Gerold', 'male'),
```

('Belicia', 'female'),
('Esma', 'female'),
('Kippy', 'female'),
('Marion', 'male'),
('Morgana', 'female'),
('Ardella', 'female'),
('Joey', 'male'),
('Irving', 'male'),
('Ricca', 'female'),
('Dryke', 'male'),
('Noah', 'male'),
('Thorsten', 'male'),
('Joellen', 'female'),
('Aeriel', 'female'),
('Fredra', 'female'),
('Marlo', 'female'),
('Ismail', 'male'),
('Stacy', 'female'),
('Gelya', 'female'),
('Aradra', 'female'),
('Reynolds', 'male'),
('Agamemnon', 'male'),
('Billie', 'female'),
('Cynthea', 'female'),
('Maxy', 'female'),
('Dodie', 'female'),
('Gael', 'female'),
('Zita', 'female'),
('Edmund', 'male'),
('Meir', 'male'),
('Robena', 'female'),
('Waiter', 'male'),
('Marty', 'male'),
('Praneetf', 'male'),
('Clarisa', 'female'),
('Irena', 'female'),
('Oriana', 'female'),
('Maribel', 'female'),
('Audra', 'female'),
('Camile', 'female'),
('Wittie', 'male'),
('Ethel', 'female'),
('Towny', 'male'),
('Florentia', 'female'),
('Ada', 'female'),
('Lelah', 'female'),
('Sonni', 'female'),
('Jannelle', 'female'),
('Lisandra', 'female'),
('Wini', 'female'),
('Natala', 'female'),
('Eddie', 'male'),
('Sheena', 'female'),
('Gypsy', 'female'),
('Cornela', 'female'),
('Douggee', 'male'),

('Jess', 'male'),
('Sullivan', 'male'),
('Jami', 'female'),
('Johannes', 'male'),
('Knox', 'male'),
('Grethel', 'female'),
('Millisent', 'female'),
('Tann', 'male'),
('Kessia', 'female'),
('Elise', 'female'),
('Luciano', 'male'),
('Junette', 'female'),
('Garwood', 'male'),
('Pat', 'male'),
('Diana', 'female'),
('Porter', 'male'),
('Leone', 'female'),
('Rudy', 'male'),
('Georgianna', 'female'),
('Hilary', 'female'),
('Taylor', 'male'),
('Vanda', 'female'),
('Aubrey', 'male'),
('Leo', 'male'),
('Sharai', 'female'),
('Dannye', 'female'),
('Adriane', 'female'),
('Gaspar', 'male'),
('Harmonie', 'female'),
('Friedrick', 'male'),
('Rosemarie', 'female'),
('Hewet', 'male'),
('Emmit', 'male'),
('Jillane', 'female'),
('Guy', 'male'),
('Timothea', 'female'),
('Rubetta', 'female'),
('Bendite', 'female'),
('Phineas', 'male'),
('Shelley', 'male'),
('Salim', 'male'),
('Mart', 'male'),
('Sharyl', 'female'),
('Doreen', 'female'),
('Annadiane', 'female'),
('Johnna', 'female'),
('Kendal', 'male'),
('Jabez', 'male'),
('Wolfy', 'male'),
('Jammie', 'female'),
('Nico', 'male'),
('Rikki', 'female'),
('Myrna', 'female'),
('Christian', 'male'),
('Neilla', 'female'),
('Golda', 'female'),

('Kellen', 'female'),
('Ingmar', 'male'),
('Philippe', 'female'),
('Ralf', 'male'),
('Harlan', 'male'),
('Alia', 'female'),
('Kory', 'male'),
('Weider', 'male'),
('Torry', 'male'),
('Anne', 'female'),
('Preston', 'male'),
('Rosamond', 'female'),
('Red', 'male'),
('Kessiah', 'female'),
('Valene', 'female'),
('Selma', 'female'),
('Lorry', 'female'),
('Auguste', 'female'),
('Gaynor', 'female'),
('Carroll', 'male'),
('Trevar', 'male'),
('Fanchette', 'female'),
('Brewster', 'male'),
('Lanni', 'female'),
('Annemarie', 'female'),
('Rockwell', 'male'),
('Hilliard', 'male'),
('Shoshie', 'female'),
('Chen', 'male'),
('Philomena', 'female'),
('Dode', 'female'),
('Rowena', 'female'),
('Don', 'male'),
('Donnamarie', 'female'),
('Devina', 'female'),
('Sandy', 'male'),
('Idette', 'female'),
('Daile', 'female'),
('Mair', 'female'),
('Alexei', 'male'),
('Kevan', 'male'),
('Ernie', 'male'),
('Annabel', 'female'),
('Agustin', 'male'),
('Deerdre', 'female'),
('Margareta', 'female'),
('Sean', 'male'),
('Stormy', 'female'),
('Laurent', 'male'),
('Peri', 'female'),
('Klarika', 'female'),
('Maxie', 'female'),
('Giancarlo', 'male'),
('Trudi', 'female'),
('Brita', 'female'),
('Korie', 'female'),

('Von', 'male'),
('Giovanni', 'male'),
('Vaughan', 'male'),
('Barnie', 'male'),
('Barnebas', 'male'),
('Aguinaldo', 'male'),
('Gretna', 'female'),
('Gabbie', 'female'),
('Dollie', 'female'),
('Jinny', 'female'),
('Nevins', 'male'),
('Nora', 'female'),
('Sallyann', 'female'),
('Edeline', 'female'),
('Leonardo', 'male'),
('Stanley', 'male'),
('Ravi', 'male'),
('Ag', 'female'),
('Loralyn', 'female'),
('Sheldon', 'male'),
('Reba', 'female'),
('Dynah', 'female'),
('Joanna', 'female'),
('Codee', 'female'),
('Anna-Diana', 'female'),
('Stacie', 'female'),
('Jimmie', 'male'),
('Eyde', 'female'),
('Amalie', 'female'),
('Vite', 'male'),
('Theodoric', 'male'),
('Liora', 'female'),
('Abel', 'male'),
('Lilah', 'female'),
('Bailie', 'male'),
('Iolanthe', 'female'),
('Bertram', 'male'),
('Lorna', 'female'),
('Jehu', 'male'),
('Partha', 'male'),
('Lynett', 'female'),
('Willetta', 'female'),
('Aubine', 'female'),
('Trina', 'female'),
('Andreana', 'female'),
('Pierson', 'male'),
('Katy', 'female'),
('Sherry', 'female'),
('Merola', 'female'),
('Angelika', 'female'),
('Aleda', 'female'),
('Nicolette', 'female'),
('Dixie', 'female'),
('Stillman', 'male'),
('Jenni', 'female'),
('Clare', 'female'),

('Parke', 'male'),
('Connor', 'male'),
('Franky', 'male'),
('Emmalee', 'female'),
('Puff', 'male'),
('Madlin', 'female'),
('Nathalie', 'female'),
('Leesa', 'female'),
('Heidi', 'female'),
('Verene', 'female'),
('Zaneta', 'female'),
('Zedekiah', 'male'),
('Mab', 'female'),
('Albert', 'male'),
('Grazia', 'female'),
('Erica', 'female'),
('Sibyl', 'female'),
('Parker', 'male'),
('Johann', 'male'),
('Gerhardine', 'female'),
('Chloris', 'female'),
('Merlin', 'male'),
('Ray', 'male'),
('Bridgette', 'female'),
('Hilde', 'female'),
('Micheal', 'male'),
('Esther', 'female'),
('Angelo', 'male'),
('Hali', 'female'),
('Janaye', 'female'),
('Elsa', 'female'),
('Felicia', 'female'),
('Noemi', 'female'),
('Bettina', 'female'),
('Dustin', 'male'),
('Tannie', 'male'),
('Saraann', 'female'),
('Celestia', 'female'),
('Brunhilda', 'female'),
('Leanna', 'female'),
('Kiley', 'female'),
('Chas', 'male'),
('Krystle', 'female'),
('Barby', 'female'),
('Ella', 'female'),
('Shel', 'female'),
('Hezekiah', 'male'),
('Codi', 'female'),
('Endora', 'female'),
('Guthrey', 'male'),
('Charisse', 'female'),
('Iggie', 'male'),
('Vonnie', 'female'),
('Vinita', 'female'),
('Ethelred', 'male'),
('Stevena', 'female'),

('Kerry', 'female'),
('Hyatt', 'male'),
('Celinda', 'female'),
('Cherilynn', 'female'),
('Dawna', 'female'),
('Betsey', 'female'),
('Joane', 'female'),
('Jesse', 'female'),
('Rufe', 'male'),
('Tudor', 'male'),
('Mallory', 'female'),
('Haily', 'female'),
('Rivy', 'female'),
('Hillary', 'female'),
('Lilia', 'female'),
('Foster', 'male'),
('Bianka', 'female'),
('Garvy', 'male'),
('Fiann', 'female'),
('Herbie', 'male'),
('Desdemona', 'female'),
('Syd', 'male'),
('Benedicta', 'female'),
('Wendeline', 'female'),
('Katinka', 'female'),
('Teddie', 'male'),
('Titus', 'male'),
('Merrilee', 'female'),
('Petr', 'male'),
('Elke', 'female'),
('Noelyn', 'female'),
('Paten', 'male'),
('Miriam', 'female'),
('Verge', 'male'),
('Aurore', 'female'),
('Melinda', 'female'),
('Leonanie', 'female'),
('Ethan', 'male'),
('Adrick', 'male'),
('Arel', 'male'),
('Raymund', 'male'),
('Britte', 'female'),
('Wit', 'male'),
('Giles', 'male'),
('Orazio', 'male'),
('Lovell', 'male'),
('Kaspar', 'male'),
('Ardath', 'female'),
('Amandi', 'female'),
('Genni', 'female'),
('Rusty', 'male'),
('Irina', 'female'),
('Aime', 'female'),
('Tybalt', 'male'),
('Terri', 'female'),
('Stanford', 'male'),

('Dominick', 'male'),
('Nils', 'male'),
('Antonino', 'male'),
('Wiatt', 'male'),
('Rahul', 'male'),
('Bartie', 'male'),
('Kettie', 'female'),
('Mindy', 'female'),
('Elena', 'female'),
('Mahalia', 'female'),
('Guido', 'male'),
('Morgan', 'male'),
('Phillipe', 'male'),
('Yevette', 'female'),
('Rosaleen', 'female'),
('Tobin', 'male'),
('Pepillo', 'male'),
('Cathy', 'female'),
('Ethelyn', 'female'),
('Prudence', 'female'),
('Mahmud', 'male'),
('Sherline', 'female'),
('Kirsteni', 'female'),
('Laureen', 'female'),
('Aleksandrs', 'male'),
('Ali', 'male'),
('Umeko', 'female'),
('Storey', 'female'),
('Webster', 'male'),
('Arlyne', 'female'),
('Dodi', 'female'),
('Laverna', 'female'),
('Iggy', 'male'),
('Diego', 'male'),
('Darsie', 'female'),
('Rosie', 'female'),
('Merralee', 'female'),
('Enrichetta', 'female'),
('Dix', 'female'),
('Alexandrina', 'female'),
('Jordon', 'male'),
('Melonie', 'female'),
('Karisa', 'female'),
('Rodney', 'male'),
('Ninnette', 'female'),
('Corabel', 'female'),
('Sabra', 'female'),
('Keenan', 'male'),
('Glynnis', 'female'),
('Ellyn', 'female'),
('Francesmary', 'female'),
('Evita', 'female'),
('Elmore', 'male'),
('Pansie', 'female'),
('Sherill', 'female'),
('Van', 'female'),

('Shimon', 'male'),
('Sal', 'female'),
('Armando', 'male'),
('Meredeth', 'male'),
('Dacey', 'female'),
('Jacki', 'female'),
('Lucita', 'female'),
('Edythe', 'female'),
('Gerald', 'male'),
('Sampson', 'male'),
('Ilene', 'female'),
('Kaster', 'female'),
('Tab', 'male'),
('Fionna', 'female'),
('Fran', 'female'),
('Phil', 'male'),
('Fina', 'female'),
('Winn', 'male'),
('Livia', 'female'),
('Margaret', 'female'),
('Oran', 'male'),
('Powell', 'male'),
('Derron', 'male'),
('Lockwood', 'male'),
('Lynnell', 'female'),
('Elsie', 'female'),
('Taryn', 'female'),
('Riley', 'male'),
('Lynelle', 'female'),
('Fraser', 'male'),
('Jackquelin', 'female'),
('Leola', 'female'),
('Ezekiel', 'male'),
('Cesya', 'female'),
('Charita', 'female'),
('Karlen', 'female'),
('Herve', 'male'),
('Ines', 'female'),
('Adams', 'male'),
('Laurene', 'female'),
('Felix', 'male'),
('Oren', 'male'),
('Nisa', 'female'),
('Alejandra', 'female'),
('Claudetta', 'female'),
('Liana', 'female'),
('Jorry', 'female'),
('Janos', 'male'),
('Tiebold', 'male'),
('Lenee', 'female'),
('Arlo', 'male'),
('Ulrike', 'female'),
('Lemuel', 'male'),
('Rebeka', 'female'),
('Leyla', 'female'),
('Collie', 'female'),

('Theresa', 'female'),
('Huey', 'male'),
('Vally', 'female'),
('Marsiella', 'female'),
('Darius', 'male'),
('Philly', 'female'),
('Albertina', 'female'),
('Osbourne', 'male'),
('Jonis', 'female'),
('Novelia', 'female'),
('Ferinand', 'male'),
('Langston', 'male'),
('Isaak', 'male'),
('Orren', 'male'),
('Alain', 'male'),
('Haley', 'female'),
('Fritz', 'male'),
('Constancy', 'female'),
('Hortense', 'female'),
('Ronni', 'female'),
('Catie', 'female'),
('Woody', 'male'),
('Brandice', 'female'),
('Jean-Francois', 'male'),
('Helga', 'female'),
('Bethina', 'female'),
('Ted', 'male'),
('Bonnie', 'female'),
('Meyer', 'male'),
('Damaris', 'female'),
('Frederick', 'male'),
('Donella', 'female'),
('Jaquenette', 'female'),
('Rudie', 'male'),
('Devi', 'female'),
('Bayard', 'male'),
('Friederike', 'female'),
('Dody', 'female'),
('Darcee', 'female'),
('Chevalier', 'male'),
('Bartel', 'male'),
('Blondelle', 'female'),
('Dean', 'male'),
('Odessa', 'female'),
('Mohamed', 'male'),
('Kelsi', 'female'),
('Willie', 'female'),
('Fairfax', 'male'),
('Nellie', 'female'),
('Jeniece', 'female'),
('Queenie', 'female'),
('Moishe', 'male'),
('Casey', 'male'),
('Ingaberg', 'female'),
('Wilbert', 'male'),
('Skylar', 'male'),

('Carmelina', 'female'),
('Stefania', 'female'),
('Kelli', 'female'),
('Rachele', 'female'),
('Tish', 'female'),
('Farica', 'female'),
('Pattie', 'female'),
('Delly', 'female'),
('Keelia', 'female'),
('Carena', 'female'),
('Dayle', 'female'),
('Andre', 'male'),
('Karole', 'female'),
('Mirabel', 'female'),
('Donica', 'female'),
('Carita', 'female'),
('Washington', 'male'),
('Siana', 'female'),
('Dorey', 'female'),
('Clifford', 'male'),
('Monte', 'male'),
('Robbi', 'female'),
('Malynda', 'female'),
('Nerte', 'female'),
('Valdemar', 'male'),
('Quintin', 'male'),
('Mallissa', 'female'),
('Lynsey', 'female'),
('Kiri', 'female'),
('Gennifer', 'female'),
('Alfonse', 'male'),
('Sile', 'female'),
('Nena', 'female'),
('Bernadine', 'female'),
('Val', 'female'),
('Vale', 'female'),
('Ignacio', 'male'),
('Cindi', 'female'),
('Karlee', 'female'),
('Wendie', 'female'),
('Fortune', 'female'),
('Regina', 'female'),
('Shaine', 'male'),
('Clareta', 'female'),
('Danika', 'female'),
('Randee', 'female'),
('Celestine', 'female'),
('Chrissie', 'female'),
('Halimeda', 'female'),
('Annalee', 'female'),
('Grover', 'male'),
('Zilvia', 'female'),
('Germaine', 'male'),
('Celene', 'female'),
('Reeta', 'female'),
('Jewelle', 'female'),

('Cornie', 'female'),
('Dimitry', 'male'),
('Diannne', 'female'),
('Werner', 'male'),
('Susi', 'female'),
('Reinhard', 'male'),
('Caresse', 'female'),
('Patty', 'male'),
('Tan', 'male'),
('Antonie', 'female'),
('Gwenore', 'female'),
('Leda', 'female'),
('Fawnia', 'female'),
('Darcie', 'female'),
('Adaline', 'female'),
('Katine', 'female'),
('Cara', 'female'),
('Giacomo', 'male'),
('Mel', 'female'),
('Fanechka', 'female'),
('Mort', 'male'),
('Freda', 'female'),
('Vita', 'female'),
('Roselin', 'female'),
('Ewan', 'male'),
('Malorie', 'female'),
('Vasily', 'male'),
('Godfry', 'male'),
('Lancelot', 'male'),
('Gustavus', 'male'),
('Milissent', 'female'),
('Vena', 'female'),
('Fionnula', 'female'),
('Gianna', 'female'),
('Jamey', 'male'),
('Chandal', 'female'),
('Benny', 'female'),
('Ed', 'male'),
('Melly', 'female'),
('Katuscha', 'female'),
('Jessy', 'female'),
('Parsifal', 'male'),
('Harriot', 'female'),
('Opal', 'female'),
('Bogart', 'male'),
('Daffy', 'female'),
('Alvin', 'male'),
('Bethena', 'female'),
('Antony', 'male'),
('Kath', 'female'),
('Cindy', 'female'),
('Carey', 'female'),
('Fayette', 'female'),
('Averil', 'male'),
('Euphemia', 'female'),
('Stacey', 'female'),

('Pierette', 'female'),
('Letisha', 'female'),
('Herta', 'female'),
('Allene', 'female'),
('Hal', 'male'),
('Godart', 'male'),
('Thorpe', 'male'),
('Lita', 'female'),
('Paddie', 'male'),
('Burgess', 'male'),
('Rochell', 'female'),
('Valaria', 'female'),
('Peyton', 'male'),
('Rog', 'male'),
('Barnaby', 'male'),
('Saw', 'male'),
('Chelsea', 'female'),
('Catlin', 'female'),
('Michael', 'male'),
('Daryl', 'male'),
('Clemens', 'male'),
('Cilka', 'female'),
('Melisa', 'female'),
('Selig', 'male'),
('Joel', 'male'),
('Flore', 'female'),
('Claude', 'female'),
('Cheslie', 'female'),
('Dimitrou', 'male'),
('Marti', 'female'),
('Mariann', 'female'),
('Lorilee', 'female'),
('Aprilette', 'female'),
('Luigi', 'male'),
('Aleece', 'female'),
('Standford', 'male'),
('Sanderson', 'male'),
('Casey', 'female'),
('Elfreda', 'female'),
('Emmy', 'male'),
('Monah', 'female'),
('Allianora', 'female'),
('Morgen', 'female'),
('Vivian', 'female'),
('Antonietta', 'female'),
('Kamillah', 'female'),
('Yank', 'male'),
('Andri', 'male'),
('Bart', 'male'),
('Ramsey', 'male'),
('Ilka', 'female'),
('Damon', 'male'),
('Jeffery', 'male'),
('Aubrie', 'female'),
('Eulalie', 'female'),
('Mira', 'female'),

('Meier', 'male'),
('Frazier', 'male'),
('Sybil', 'female'),
('Edna', 'female'),
('Kiah', 'female'),
('Gilburt', 'male'),
('Marena', 'female'),
('Emyle', 'female'),
('Doris', 'female'),
('Kendall', 'male'),
('Malkah', 'female'),
('Archibold', 'male'),
('Violette', 'female'),
('Cherice', 'female'),
('Leopold', 'male'),
('Marcio', 'male'),
('Pandora', 'female'),
('Dwight', 'male'),
('Karel', 'male'),
('Eldon', 'male'),
('Melina', 'female'),
('Rosamund', 'female'),
('Valentina', 'female'),
('Hewett', 'male'),
('Tandy', 'female'),
('Lucienne', 'female'),
('Melisent', 'female'),
('Ollie', 'male'),
('Ike', 'female'),
('Olwen', 'female'),
('West', 'male'),
('Agata', 'female'),
('Moise', 'male'),
('Trip', 'male'),
('Sammie', 'male'),
('Charmian', 'female'),
('Glyn', 'female'),
('Melba', 'female'),
('Chere', 'female'),
('Niki', 'male'),
('Jakob', 'male'),
('Bradly', 'male'),
('Darlene', 'female'),
('Louisette', 'female'),
('Ann-Mari', 'female'),
('Madge', 'female'),
('Welby', 'male'),
('Sargent', 'male'),
('Sanford', 'male'),
('Horace', 'male'),
('Elisha', 'male'),
('Rene', 'male'),
('Britney', 'female'),
('Dyane', 'female'),
('Anselm', 'male'),
('Derek', 'male'),

('Willis', 'male'),
('Ainslee', 'female'),
('Trudie', 'female'),
('Lyle', 'male'),
('Katrina', 'female'),
('Daisy', 'female'),
('Cal', 'female'),
('Alanna', 'female'),
('Catlee', 'female'),
('Stephie', 'female'),
('Wolfram', 'male'),
('Brant', 'male'),
('Christy', 'male'),
('Isador', 'female'),
('Torrance', 'male'),
('Ceciley', 'female'),
('Robbie', 'female'),
('Tiphannie', 'female'),
('Ricky', 'male'),
('Eberhard', 'male'),
('Adolpho', 'male'),
('Nanon', 'female'),
('Joe', 'male'),
('Andi', 'female'),
('Abby', 'male'),
('Dena', 'female'),
('Nick', 'male'),
('Nina', 'female'),
('Marcile', 'female'),
('Ami', 'female'),
('Lyndsey', 'female'),
('Charlott', 'female'),
('Tye', 'male'),
('Alexi', 'female'),
('Halli', 'female'),
('Levy', 'male'),
('Marcel', 'male'),
('Suzanna', 'female'),
('Abelard', 'male'),
('Lesli', 'female'),
('Vladimir', 'male'),
('Lindsy', 'female'),
('Lamar', 'male'),
('Dorette', 'female'),
('Adrian', 'male'),
('Koren', 'female'),
('Cathe', 'female'),
('Rochella', 'female'),
('Florri', 'female'),
('Antoni', 'male'),
('Pinchas', 'male'),
('Bartlett', 'male'),
('Monty', 'male'),
('Saxe', 'male'),
('Teri', 'female'),
('Deborah', 'female'),

('Wenona', 'female'),
('Sibel', 'female'),
('Reine', 'female'),
('Beverlie', 'female'),
('Amos', 'male'),
('Alys', 'female'),
('Lela', 'female'),
('Bridgett', 'female'),
('Bishop', 'male'),
('Valenka', 'female'),
('Liam', 'male'),
('Lem', 'male'),
('Woodie', 'male'),
('Elwira', 'female'),
('Rodrique', 'male'),
('Prentice', 'male'),
('Leontine', 'female'),
('Englebart', 'male'),
('Marlo', 'male'),
('Miles', 'male'),
('Jethro', 'male'),
('Cordie', 'female'),
('Charlotte', 'female'),
('Marketa', 'female'),
('Helaina', 'female'),
('Dorolice', 'female'),
('Cher', 'female'),
('Imogen', 'female'),
('Muffin', 'male'),
('Hervey', 'male'),
('Pembroke', 'male'),
('Pascal', 'male'),
('Steward', 'male'),
('Worthington', 'male'),
('Franklin', 'male'),
('Jodie', 'female'),
('Zollie', 'male'),
('Konstanze', 'female'),
('Carma', 'female'),
('Korney', 'female'),
('Gaylene', 'female'),
('Erwin', 'male'),
('Sayer', 'male'),
('Shaylah', 'female'),
('Remy', 'female'),
('Brandy', 'male'),
('Matelda', 'female'),
('Tamas', 'male'),
('Melisande', 'female'),
('Clem', 'male'),
('Linda', 'female'),
('Brietta', 'female'),
('Cayla', 'female'),
('Morgan', 'female'),
('Yankee', 'male'),
('Wilow', 'female'),

('Alana', 'female'),
('Benito', 'male'),
('Elly', 'female'),
('Halvard', 'male'),
('Barbabas', 'male'),
('Suellen', 'female'),
('Cheri', 'female'),
('Orelia', 'female'),
('Elfie', 'female'),
('Bea', 'female'),
('Henka', 'female'),
('Heddie', 'female'),
('Alfreda', 'female'),
('Sissy', 'female'),
('Gay', 'male'),
('Rania', 'female'),
('Howie', 'male'),
('Clemmie', 'male'),
('Engelbart', 'male'),
('Simonette', 'female'),
('Emalia', 'female'),
('Shelbi', 'female'),
('Rodie', 'female'),
('Lisa', 'female'),
('Lane', 'female'),
('Sianna', 'female'),
('Ernesta', 'female'),
('Miguelita', 'female'),
('Lora', 'female'),
('Corrienne', 'female'),
('Marietta', 'female'),
('Matty', 'male'),
('Salmon', 'male'),
('Ulla', 'female'),
('Teane', 'female'),
('Chan', 'male'),
('Tommy', 'male'),
('Jobyna', 'female'),
('Tierney', 'female'),
('Willy', 'female'),
('Stephenie', 'female'),
('Sebastien', 'male'),
('Jolee', 'female'),
('Gerri', 'male'),
('Garnette', 'female'),
('Loree', 'female'),
('Edie', 'male'),
('Blithe', 'female'),
('Nealy', 'male'),
('Evelyn', 'male'),
('Ethelda', 'female'),
('Pooh', 'male'),
('Geof', 'male'),
('Tommy', 'female'),
('Jarvis', 'male'),
('Roslyn', 'female'),

```

('Lilla', 'female'),
('Holly', 'female'),
('Louie', 'male'),
('Gerome', 'male'),
('Annalisa', 'female'),
('Charlot', 'female'),
('Clayborn', 'male'),
('Myrtie', 'female'),
('Margalit', 'female'),
('Steffen', 'male'),
('Merrielle', 'female'),
('Kali', 'female'),
('Jameson', 'male'),
('Leanora', 'female'),
('Willard', 'male'),
('Letta', 'female'),
('Ellen', 'female'),
('Sanders', 'male'),
('Gaby', 'female'),
('Denice', 'female'),
('Brice', 'male'),
('Secunda', 'female'),
('Evonne', 'female'),
('Ronny', 'female'),
('Ilysa', 'female'),
('Nalani', 'female'),
('Austen', 'male'),
('Blythe', 'female'),
('Kala', 'female'),
('Felita', 'female'),
('Muire', 'female'),
('Luci', 'female'),
('Randi', 'female'),
('Lisabeth', 'female'),
('Tonnie', 'male'),
('Maybelle', 'female'),
('Nolan', 'male'),
('Lemmy', 'male'),
('Ariella', 'female'),
('Ibbie', 'female'),
('Stillmann', 'male'),
('Fay', 'female'),
('Alec', 'male'),
('Amalia', 'female'),
('Deina', 'female'),
('Andrei', 'female'),
('Janean', 'female'),
('Fawna', 'female'),
...]
```

```

In [12]: from sklearn.feature_extraction.text import CountVectorizer
vect = CountVectorizer(binary =True) # Explore other parameters too!!
corpus = ["Tesseract is good optical character recognition engine ", "optical
vect.fit(corpus) #learns the unique words (Vocab) in the corpus
```

```

#vect.vocabulary_ #stores the vocab #vocabulary_ built-in:-)
vocab = vect.vocabulary_

for key in sorted(vocab.keys()):
    print("{}:{}".format(key, vocab[key]))

print(vect.transform(["This is a good optical illusion"]).toarray())

#Finding Similarity Between the Documents
from sklearn.metrics.pairwise import cosine_similarity
similarity = cosine_similarity(vect.transform(["Google Cloud Vision is a cha
print(similarity)

```

```

character:0
engine:1
good:2
is:3
optical:4
recognition:5
significant:6
tesseract:7
[[0 0 1 1 1 0 0 0]]
[[0.89442719]]

```

```

In [13]: from nltk.corpus import names
names.words()
print(len(names.words()))

labeled_names = [(name, 'male') for name in names.words('male.txt')] + [(name, 'female') for name in names.words('female.txt')]

import random
random.shuffle(labeled_names)

featuresets = [(gender_features(n), gender) for (n, gender) in labeled_names]
#Try Printing featuresets

train_set, test_set = featuresets[5000:], featuresets[:2000]

import nltk
classifier = nltk.NaiveBayesClassifier.train(train_set)

classifier.classify(gender_features('Obama'))
#classifier.classify(gender_features('Yann Lecun'))

#print(nltk.classify.accuracy(classifier, test_set))

```

```
7944
```

```
Out[13]: 'female'
```

```
In [ ]:
```