Using average vector as the centroids

Microsoft Windows [Version 6.3.9600]

(c) 2013 Microsoft Corporation. All rights reserved.

C:\Users\Alex>cd Desktop

C:\Users\Alex\Desktop>python kmeans.py

Processing dataset...

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 2, iterations 200

Macro Averaged Precision over

0.182000990891

Macro Averaged Recall over

0.677352941176

Macro Averaged F Score over

0.286392570374

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 3, iterations 200

Macro Averaged Precision over

0.209208211585

Macro Averaged Recall over

0.523333333333

Macro Averaged F Score over

0.297211019397

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 4, iterations 200

Macro Averaged Precision over

0.230261581163

Macro Averaged Recall over

0.446037581699

Macro Averaged F Score over

0.302255509288

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 5, iterations 200

Macro Averaged Precision over

0.246113696435

Macro Averaged Recall over

0.406612745098

Macro Averaged F Score over

0.304457527232

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 6, iterations 200

Macro Averaged Precision over

0.254861038755

Macro Averaged Recall over

0.376844771242

Macro Averaged F Score over

0.301891689707

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 7, iterations 200

Macro Averaged Precision over

0.268634228155

Macro Averaged Recall over

0.356597805789

Macro Averaged F Score over

0.303844475725

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 8, iterations 200

Macro Averaged Precision over

0.276261909523

Macro Averaged Recall over

0.341621498599

Macro Averaged F Score over

0.303463183125

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 9, iterations 200

Macro Averaged Precision over

0.285294106106

Macro Averaged Recall over

0.341216503268

Macro Averaged F Score over

0.309025533261

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 10, iterations 200

Macro Averaged Precision over

0.292963968242

Macro Averaged Recall over

0.334348039216

Macro Averaged F Score over

0.310585751898

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 11, iterations 200

Macro Averaged Precision over

0.298377208909

Macro Averaged Recall over

0.332489262372

Macro Averaged F Score over

0.312766743597

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 12, iterations 200

Macro Averaged Precision over

0.306923272886

Macro Averaged Recall over

0.326295401494

Macro Averaged F Score over

0.315038556681

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 13, iterations 200

Macro Averaged Precision over

0.313480115954

Macro Averaged Recall over

0.329691876751

Macro Averaged F Score over

0.319675311572

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 14, iterations 200

Macro Averaged Precision over

0.318397264873

Macro Averaged Recall over

0.326287931839

Macro Averaged F Score over

0.321107061275

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 15, iterations 200

Macro Averaged Precision over

0.324452721722

Macro Averaged Recall over

0.329030812325

Macro Averaged F Score over

0.325469537488

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 16, iterations 200

Macro Averaged Precision over

0.327294465622

Macro Averaged Recall over

0.334116479925

Macro Averaged F Score over

0.329560490996

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 17, iterations 200

Macro Averaged Precision over

0.335583807246

Macro Averaged Recall over

0.332192343604

Macro Averaged F Score over

0.333004129078

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 18, iterations 200

Macro Averaged Precision over

0.339798225803

Macro Averaged Recall over

0.334598506069

Macro Averaged F Score over

0.336373026387

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 19, iterations 200

Macro Averaged Precision over

0.34355715836

Macro Averaged Recall over

0.331629318394

Macro Averaged F Score over

0.336880725884

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 20, iterations 200

Macro Averaged Precision over

0.347065159954

Macro Averaged Recall over

0.338964752568

Macro Averaged F Score over

0.342327475832

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

C:\Users\Alex\Desktop>

Using nearest datapoint to average vector as the centroids

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 2, iterations 200

Macro Averaged Precision over

0.18059373536

Macro Averaged Recall over

0.668676470588

Macro Averaged F Score over

0.283910526525

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 3, iterations 200

Macro Averaged Precision over

0.208283739148

Macro Averaged Recall over

0.525718954248

Macro Averaged F Score over

0.29721165195

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 4, iterations 200

Macro Averaged Precision over

0.226801321428

Macro Averaged Recall over

0.44

Macro Averaged F Score over

0.297977062316

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 5, iterations 200

Macro Averaged Precision over

0.245591274109

Macro Averaged Recall over

0.406637254902

Macro Averaged F Score over

0.304297553515

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 6, iterations 200

Macro Averaged Precision over

0.253883891052

Macro Averaged Recall over

0.37812254902

Macro Averaged F Score over

0.301521740572

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 7, iterations 200

Macro Averaged Precision over

0.268749825322

Macro Averaged Recall over

0. 35764729225

Macro Averaged F Score over

0.304753206584

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 8, iterations 200

Macro Averaged Precision over

0.276343659893

Macro Averaged Recall over

0.342384920635

Macro Averaged F Score over

0.303731093227

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 9, iterations 200

Macro Averaged Precision over

0.288726218339

Macro Averaged Recall over

0.342472572362

Macro Averaged F Score over

0.311156593717

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 10, iterations 200

Macro Averaged Precision over

0.293085511122

Macro Averaged Recall over

0.335302521008

Macro Averaged F Score over

0.310729780844

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 11, iterations 200

Macro Averaged Precision over

0.303406431268

Macro Averaged Recall over

0.327029878618

Macro Averaged F Score over

0.313008041248

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 12, iterations 200

Macro Averaged Precision over

0.309193309643

Macro Averaged Recall over

0.32764869281

Macro Averaged F Score over

0.316672847713

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 13, iterations 200

Macro Averaged Precision over

0.310336057167

Macro Averaged Recall over

0.329840802988

Macro Averaged F Score over

0.318336246048

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 14, iterations 200

Macro Averaged Precision over

0.318929094371

Macro Averaged Recall over

0.323554038282

Macro Averaged F Score over

0.31989103763

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 15, iterations 200

Macro Averaged Precision over

0.324229492366

Macro Averaged Recall over

0.327408730159

Macro Averaged F Score over

0.32481043744

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 16, iterations 200

Macro Averaged Precision over

0.329461160276

Macro Averaged Recall over

0.328164098973

Macro Averaged F Score over

0.327872502578

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 17, iterations 200

Macro Averaged Precision over

0.335997669493

Macro Averaged Recall over

0.332691409897

Macro Averaged F Score over

0.333319007859

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 18, iterations 200

Macro Averaged Precision over

0.338261458078

Macro Averaged Recall over

0.330533963585

Macro Averaged F Score over

0.333674788993

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

k value 19, iterations 200

Macro Averaged Precision over

0.341659497553

Macro Averaged Recall over

0.339494047619

Macro Averaged F Score over

0.339788633896

k value 20, iterations 200

Macro Averaged Precision over

0.344213733268

Macro Averaged Recall over

0.33868697479

Macro Averaged F Score over

0.340872252063