

24 ΔΕΚ 2022

Παρουσίαση Project

Αναπαράσταση
Γνώσης στον
Παγκόσμιο Ιστό

Παπαδάκης Νικόλαος

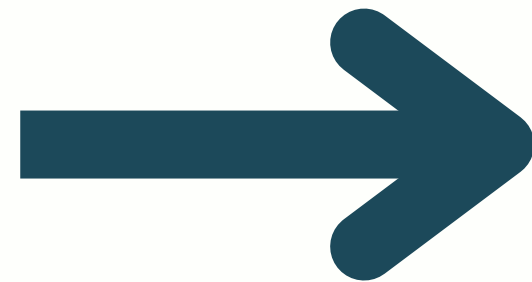


Ψαλτάκης Γιώργος ΤΗ20027



Υλοποίηση SPARQL

Χρησιμοποιώ για πρόθεμα του χώρου ονομάτων rdf ως `<http://www.w3.org/1999/02/22-rdf-syntax-ns#>`. Αυτό το πρόθεμα χρησιμοποιείται για να αναφερθεί στο χώρο ονομάτων RDF



`?type`, `?instance`, `?property`, `?value` και `?name`. Η μεταβλητή `?type` αντιπροσωπεύει τον τύπο της περίπτωσης, η μεταβλητή `?instance` αντιπροσωπεύει την ίδια την περίπτωση, η μεταβλητή `?property` αντιπροσωπεύει μια ιδιότητα της περίπτωσης, η μεταβλητή `?value` αντιπροσωπεύει την τιμή της ιδιότητας και η μεταβλητή `?name` αντιπροσωπεύει το όνομα της περίπτωσης.

Τρία τριπλά μοτίβα.

Το πρώτο τριπλό μοτίβο `?instance rdf:type ?type` ταιριάζει με κάθε τριπλό στο γράφημα όπου το υποκείμενο είναι μια περίπτωση και το αντικείμενο είναι ένας τύπος, και συνδέει το υποκείμενο με τη μεταβλητή `?instance` και το αντικείμενο με τη μεταβλητή `?type`.

Κώδικας

```
from rdflib import Graph, URIRef
from rdflib.namespace import RDF
from sklearn.feature_extraction import DictVectorizer
from sklearn.metrics.pairwise import cosine_similarity
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt

g = Graph()
g.parse("kenza_conference.rdf", format="nt")
query = """
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
SELECT ?type ?instance ?property ?value ?name
WHERE {
    ?instance rdf:type ?type .
    ?instance ?property ?value .
    OPTIONAL { ?instance <http://xmlns.com/foaf/0.1/name> ?
name }
}
"""

results = g.query(query)
instances = []
for row in results:
    instance = row[1]
    property = row[2]
    value = row[3]
    name = row[4]

    instance_dict = next((d for d in instances if d["uri"] == instance),
None)
    if instance_dict is None:
        instance_dict = {"uri": instance}
        instances.append(instance_dict)
    instance_dict[property] = value
    if name:
        instance_dict["name"] = name
```

```
vectorizer = DictVectorizer()
X = vectorizer.fit_transform(instances)
similarity_matrix = cosine_similarity(X)

sse = []

for k in range(2, 30):
    kmeans = KMeans(n_clusters=k)
    kmeans.fit(similarity_matrix)
    sse_k = kmeans.inertia_
    sse.append(sse_k)

plt.plot(range(2, 30), sse)
plt.xlabel("Number of clusters (k)")
plt.ylabel("Sum of squared errors (SSE)")
plt.show()

kmeans = KMeans(n_clusters=17)
kmeans.fit(similarity_matrix)

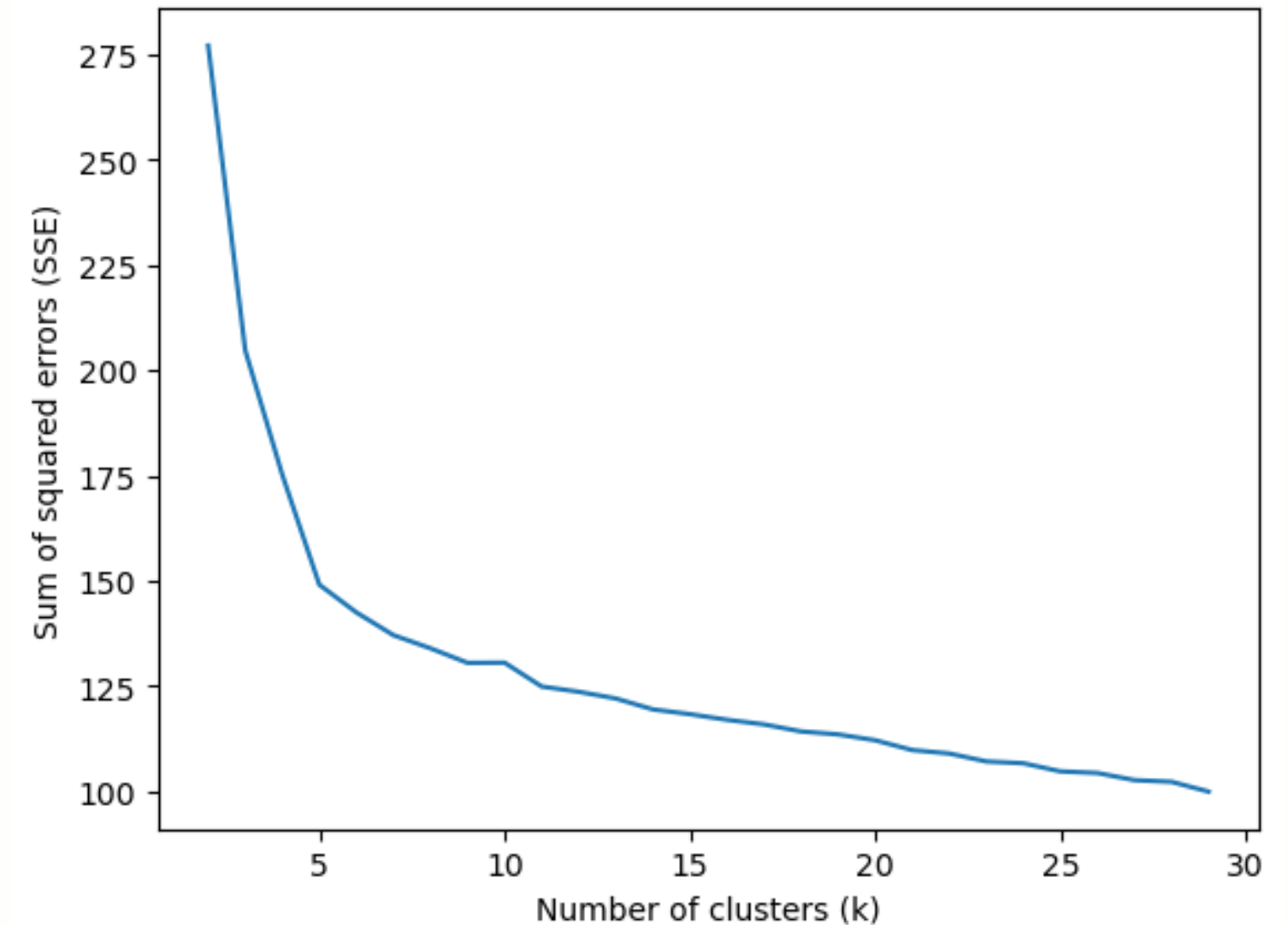
print("New types:")
print('\n')
for label in range(kmeans.n_clusters):
    cluster_instances = [instances[i] for i, l in enumerate(kmeans.labels_) if
l == label]
    print(f"Type: New type {label}")
    print('\n')
    for instance in cluster_instances:
        print(f" Instance: {instance['uri']}")
        print('\n')
        print(" Properties:")
        print('\n')
        for property, value in instance.items():
            print(f" {property}: {value}")
        print('\n')
```

Κώδικας

```
for i, instance in enumerate.instances):
    instance["cluster"] = kmeans.labels_[i]
    print(f"Instance {instance['uri']} is in cluster {instance['cluster']}")
    print('\n')
    print("Properties:")
    print('\n')
    for property, value in instance.items():
        print(f"    {property}: {value}")
    print('\n')

types = {}
for instance in instances:
    type = instance.get("type")
    if type:
        if type not in types:
            types[type] = []
        types[type].append(instance)
```

Υλοποιώ Kmeans βάση των προηγούμενων κατηγοριοποιήσεων βάση. type για να ορίσω νέα types μέσω μηχανικής μάθησης. Έτσι ορίζονται σε clusters τα νέα types απο τον γράφο με τις τριπλέτες.



Elbow

Αποτελέσματα Κατηγοριοποίησης

Type: New type 4

Instance: <http://bnode2.com>

Properties:

uri: <http://bnode2.com>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>:
http://www.w3.org/2003/01/geo/wgs84_pos#Point
http://www.w3.org/2003/01/geo/wgs84_pos#long: -80.00
http://www.w3.org/2003/01/geo/wgs84_pos#lat: 40.44

Instance: <http://bnode1.com>

Properties:

uri: <http://bnode1.com>
http://www.w3.org/2003/01/geo/wgs84_pos#lat: 40.44
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>:
http://www.w3.org/2003/01/geo/wgs84_pos#Point
http://www.w3.org/2003/01/geo/wgs84_pos#long: -80.00

Αποτελέσματα Κατηγοριοποίησης

Type: New type 7

Instance: <http://data.semanticweb.org/person/mitsuharu-nagamori>

Properties:

uri: <http://data.semanticweb.org/person/mitsuharu-nagamori><http://swrc.ontoware.org/ontology#affiliation:>
<http://data.semanticweb.org/organization/university-of-tsukuba>
name: Mitsuharu Nagamori
[http://xmlns.com/foaf/0.1/name: Mitsuharu Nagamori](http://xmlns.com/foaf/0.1/name:MitsuharuNagamori)
[http://www.w3.org/2000/01/rdf-schema#label: Mitsuharu Nagamori](http://www.w3.org/2000/01/rdf-schema#label:MitsuharuNagamori)
[http://xmlns.com/foaf/0.1/based_near: http://dbpedia.org/resource/Japan](http://xmlns.com/foaf/0.1/based_near:http://dbpedia.org/resource/Japan)<http://www.w3.org/1999/02/22-rdf-syntax-ns#type:>
<http://xmlns.com/foaf/0.1/Person>[http://data.semanticweb.org/ns/swc/ontology#holdsRole: http://data.semanticweb.org/conference/dc/2010/programme-committee-member](http://data.semanticweb.org/ns/swc/ontology#holdsRole:)

Instance: <http://data.semanticweb.org/person/akira-miyazawa>

Properties:

uri: <http://data.semanticweb.org/person/akira-miyazawa><http://data.semanticweb.org/ns/swc/ontology#holdsRole:>
<http://data.semanticweb.org/conference/dc/2010/programme-committee-member>
name: Akira Miyazawa
[http://www.w3.org/2000/01/rdf-schema#label: Akira Miyazawa](http://www.w3.org/2000/01/rdf-schema#label:AkiraMiyazawa)
[http://www.w3.org/1999/02/22-rdf-syntax-ns#type: http://xmlns.com/foaf/0.1/Person](http://www.w3.org/1999/02/22-rdf-syntax-ns#type:http://xmlns.com/foaf/0.1/Person)<http://swrc.ontoware.org/ontology#affiliation:>
<http://data.semanticweb.org/organization/national-institute-of-informatics-japan>[http://xmlns.com/foaf/0.1/name: Akira Miyazawa](http://xmlns.com/foaf/0.1/name:AkiraMiyazawa)
[http://xmlns.com/foaf/0.1/based_near: http://dbpedia.org/resource/Japan](http://xmlns.com/foaf/0.1/based_near:http://dbpedia.org/resource/Japan)

Αποτελέσματα Κατηγοριοποίησης

Instance: <http://data.semanticweb.org/person/masahide-kanzaki>

Properties:

uri: <http://data.semanticweb.org/person/masahide-kanzaki>
name: Masahide Kanzaki
<http://xmlns.com/foaf/0.1/name: Masahide Kanzaki>
<http://swrc.ontoware.org/ontology#affiliation: http://data.semanticweb.org/organization/zenon-ltd-partners>
Masahide Kanzaki
<http://data.semanticweb.org/ns/swc/ontology#holdsRole: http://data.semanticweb.org/conference/dc/2010/programme-committee-member>
http://xmlns.com/foaf/0.1/based_near: http://dbpedia.org/resource/Japan

Instance: <http://data.semanticweb.org/person/shigeo-sugimoto>

Properties:

uri: <http://data.semanticweb.org/person/shigeo-sugimoto>
name: Shigeo Sugimoto
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type: http://xmlns.com/foaf/0.1/Person>
<http://swrc.ontoware.org/ontology#affiliation: http://data.semanticweb.org/organization/university-of-tsukuba>
<http://data.semanticweb.org/ns/swc/ontology#holdsRole: http://data.semanticweb.org/conference/dc/2010/programme-committee-member>
http://xmlns.com/foaf/0.1/based_near: http://dbpedia.org/resource/Japan
<http://xmlns.com/foaf/0.1/name: Shigeo Sugimoto>

Συμπεράσματα Κατηγοριοποίησης

Βλεπουμε οτι δημιουργεί αρκετά νέα types και τα υποκατηγοριοποιεί πιο αναλυτικά και για πιο πολλές ομάδες βάση ομοιοτήτων. Αυτό συμβαίνει για όλους τους αριθμούς clusters που έχουμε βάλει μιας και οι clusters αποτελούν τα νέα types που τα instances κατηγοριοποιούνται.

Ευχαριστώ για την παρακολούθηση.