

測試方式取回傳的電影清單前 5 名(採用 precision 來比較)

query1: “joker chaos ”

- TF-IDF model
 1. **The Dark Knight**
 2. Jaws
 3. The Dark Knight Rises
 4. **Joker**
 5. The Shawshank Redemption
- Word2Vec model
 1. Cowboy Bebop: Tengoku no tobira
 2. The Dark Knight Rises
 3. **The Dark Knight**
 4. Serbuan maut 2: Berandal
 5. High Noon
- Doc2vec model
 1. **The Dark Knight**
 2. Batman Begins
 3. **Joker**
 4. The Dark Knight Rises
 5. Dirty Harry

precision 比較:

	TF-IDF	Word2Vec	Doc2vec
precision	0.4	0.2	0.4

query2: “time machine ”

- TF-IDF model
 1. Glengarry Glen Ross
 2. Papurika
 3. **The Terminator**
 4. Wonder
 5. The Irishman
- Word2Vec model

1. Andrei Rublev
2. Harry Potter and the Deathly Hallows: Part 1
3. The Message
4. **Edge of Tomorrow**
5. Wonder

- Doc2vec model
 1. **The Terminator**
 2. Harry Potter and the Deathly Hallows: Part 1
 3. Mission: Impossible - Fallout
 4. 1917
 5. The Fifth Element

precision 比較:

	TF-IDF	Word2Vec	Doc2vec
precision	0.2	0.2	0.2

query3: “ war ”

- TF-IDF model
 1. **The Great Escape**
 2. **Lawrence of Arabia**
 3. **Hotaru no haka**
 4. **Empire of the Sun**
 5. The Searchers
- Word2Vec model
 1. **The Great Escape**
 2. **Hacksaw Ridge**
 3. **All Quiet on the Western Front**
 4. **Patton**
 5. **The Pianist**
- Doc2vec model
 1. **Patton**
 2. **Kurenai no buta**
 3. **The Pianist**
 4. **Zwartboek**
 5. **Kaze tachinu**

precision 比較:

	TF-IDF	Word2Vec	Doc2vec
precision	0.8	1.0	1.0

query4: “ Romance ”

- TF-IDF model
 1. **Call Me by Your Name**
 2. **The Apartment**
 3. **Gone with the Wind**
 4. Charade
 5. Cabaret
- Word2Vec model
return null
- Doc2vec model
 1. Cabaret
 2. **Gone with the Wind**
 3. Sweet Smell of Success
 4. **To Have and Have Not**
 5. **The Apartment**

precision 比較:

	TF-IDF	Word2Vec	Doc2vec
precision	0.6	0.0	0.6

query5: “ high school love story”

- TF-IDF model
 1. **Dazed and Confused**
 2. Ferris Bueller's Day Off
 3. Remember the Titans

4. **The Breakfast Club**
5. **Druk**

- Word2Vec model
 1. **Dazed and Confused**
 2. Die Welle
 3. **The Breakfast Club**
 4. Harry Potter and the Sorcerer's Stone
 5. Ferris Bueller's Day Off
- Doc2vec model
 1. Veer-Zaara
 2. Remember the Titans
 3. **Dazed and Confused**
 4. **Mimi wo sumaseba**
 5. **Rushmore**

precision 比較:

	TF-IDF	Word2Vec	Doc2vec
precision	0.6	0.4	0.6

Over all precision

	TF-IDF	Word2Vec	Doc2vec
precision	52%	45%	56%

Hyperparameter analysis

Word2Vec

	window=10	window=5	window=1
vector_size=50	0.36	0.32	0.24
vector_size=100	0.36	0.40	0.28

vector_size=200	0.36	0.36	0.24
-----------------	------	------	------

Doc2Vec

	window=10	window=5	window=1
vector_size=50	0.56	0.56	0.56
vector_size=100	0.68	0.64	0.64
vector_size=200	0.6	0.6	0.6