

UNIVERSITY OF ABERDEEN

CS551H: Natural Language Generation

Assessment 1

Actor's Career over Time

Submitted by

Avinash Vinayak Bagul

51987820

Table of Contents

1.1	Overview	3
1.2	Structure	3
1.1.1	Basic introductory information:	3
1.1.2	Brief information about actor's overall career life:	3
1.1.3	Decade wise information:	3
1.3	Features	4
1.4	Scope for improvement	5

1.1 Overview

The main goal of assessment one is to create a robust Natural Language Generation system which can create a brief biographic summary for given actors data. The summary includes general information, career description per each decade and highlighting any change in the career (i.e. from acting to directing or producing). Arria Studio is used for the development of this system, which is an advance NLG software with cutting edge user interface.

Data is uploaded in JSON format, and then it is extracted using objects associated with the values in the input file. The provided file contains the personal details along with the movie data for each actor. Several variables and functions were created to fulfil the requirements set in assessment.

1.2 Structure

1.1.1 Basic introductory information:

Output Text:

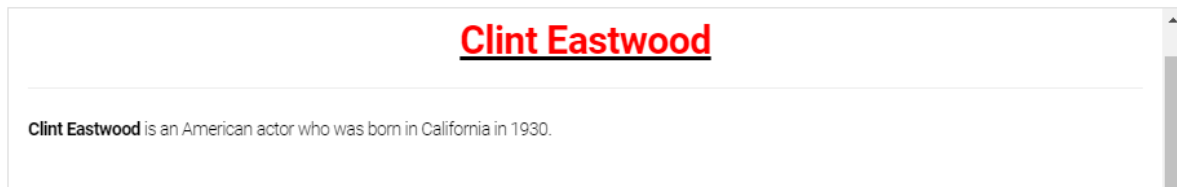


Fig (1.1.1)

Includes name, nationality, birthplace, birth year and death year (if died).

1.1.2 Brief information about actor's overall career life:

Eastwood started his career as actor in the movie "Ambush at Cimarron Pass" and "Lafayette Escadrille" in the year 1958. He acted in 52 films but did not win any academy award. Eastwood also directed 16 movies and won 2 awards. Through out his career Eastwood produced 6 movies including "The Stars Fell on Henrietta","American Sniper",etc. One of the best movie by Eastwood is "Sad Hill Unearthed" with the total rating of 8.9.

Fig (1.1.2)

Gives a descriptive summary of whole career with start of career (year and movies) total number of films as actor, producer or director with the number of awards won for the same. Also includes the best movie with highest rating from their acting career.

1.1.3 Decade wise information:

He appeared in 17 films in the 1990s., He was nominated for 2 movies and did not win any academy award. In this decade Washington's career grew at an excellent pace by acting in 13 more movies than the last decade .

He appeared in 12 films in the 2000s., He was nominated for 1 film and won one academy award. There was drop in Denzel Washington's career in this decade..

He acted in 11 movies in the 2010s., He was nominated for 3 films and did not win any academy award. There was no significant change in the career compared to previous decade .

Fig (1.1.3)

Decade summary shows total number of movies as actor in that particular decade along with the number of nominations and awards won. It also represent the growth, drop or consistency in career during the decade. If there is any change in career like in case of Clint Eastwood, it is shows the change from actor to director or producer.

1.3 Features

- **Career progression in acting, directing and producing.**

The brief summary and decade summary narrates about the career progression.

- **Random Text**

chooseAtRandom function was used to provide alternative sentences for the same context in the summary to make the summary sound more natural. It was used for some words as well as for some sentences. Mainly used in the decade summary to provide different kind of sentences in the description.

- **Change in career.**

If there is any change in career like in case of Clint Eastwood, the career change is clearly mentioned in decade where the change occurs.

In the decade 1960's, Eastwood did seven movies with no nomination and no award. In this decade Eastwood's career improved at an excellent pace by acting in 5 more movies than the last decade .

In the decade 1970's, Eastwood did 15 movies with no nomination and no award. In this decade Eastwood's career hiked at an excellent pace by acting in 8 more movies than the last decade .

In the decade 1980's, Eastwood did 12 films with no nomination and no award. In comparison with last decade acting career dropped in 1980's. .

In the decade 1990's, Eastwood did nine films with one nomination and no award. In comparison with last decade acting career dropped in 1990's. .

He acted in four movies in the 2000s., He was nominated for 1 movie and did not win any academy award. There was drop in Clint Eastwood's career in this decade. Career mainly shifted from Acting to Directing.

In the decade 2010's, Eastwood did two films with no nomination and no award. There was no significant change in the career compared to previous decade .

- **Can handle every actors data efficiently**

All the JSON files provided were tested with the system. System handles each actors efficiently and prints useful and desired information.

- **Prints only useful information.**

Only useful information is displayed for every actor. For example if someone just did acting throughout the career then his summary will not talk about directing and producing.

- **Implementation of useful functions**

Many functions were created for boosting the efficiency of system in extracting specific data. Some of the functions were to extract the details of most rated movie, number of awards, nominations, directing movies, producer movies, etc.

1.4 Scope for improvement

There is a big scope for improvement in this system. There is no limit to improving the system.

- The major drawback of the system is that, in decade summary where actor is not nominated and not awarded is still being printed. This can be easily eliminated using if-else condition, but as mentioned earlier the more we work on it the better it gets.
- Sometimes the syntax while creating a sentence gets really complex and very hard to debug. To avoid mistakes and make updating and debugging easy we can create functions or subscripts to provide us modularity and efficiency.
- Sentences could be more organised and professional based on correct English grammar.
- I personally did not use gender function to determine the gender, though I practiced it in the practical session. One improvement can be made by the use of gender function.
- Arria can be developed more for providing complex functionalities like list operations (append, remove, etc.) and so many other functionalities which can take this software on another level.