

- Aditya Agrawal : adityaai@bu.edu
- Melissa Zhen : mzhen001@bu.edu
- Liangzhuo Zhang : zhanglz@bu.edu
- Cindy Lu : cindy19@bu.edu

What is the project focus/overall goal?

1. The goal of this project is to identify inventors by race and to create an algorithm that improves Chinese and Korean character recognition.
2. Using that algorithm, then perform gender analysis of inventors based on Chinese & Korean characters.

What type of data will you collect or be analyzing?

1. We will analyze the data extracted by the previous team which contains detailed information about patents applications to USPTO from 2005 to 2022.
2. We will be collecting data from US patents and their equivalent patents in either CN or KR to extract inventor names in Korean/Chinese Characters and analyzing the differences in gender-matching.

What are potential limitations of the project?

1. There might not be enough data to make an overall accurate analysis and additional patent information may need to be pulled.
2. There are some Chinese and Korean characters that are not as common and might be missed.
3. The algorithm may be time consuming or expensive.

Why is this project important or Why did you choose this project?

1. The algorithm developed in this project will be very useful for predicting the gender of inventors with more accuracy.
2. The Chinese & Korean Character Recognition that will be developed can also be used for other tasks (like Optical Character Recognition,).
3. This project will provide a valid method for other fields like East Asian Cultural studies.
4. This project is important for AI developments (speech recognition, Computer Vision Language Translation, etc.)

What are your team's next steps? (include action items/tasks)

1. Review the previous team's work and look at the created database.
2. Match patents with both a US patent and an equivalent patent filed in at least one of the Korean Patent Office or Chinese Patent Office (like <https://patents.google.com/patent/KR101541682B1/en?q=korean+skincare&oq=korean+skincare>)
3. Gender-match inventor names using Korean and Chinese characters

4. Treating the character name as the correct match, update the previous database's work and identify different outcomes

Early Insights Presentation : Nov 07th