

Exercise 5

Submitted by Freddy Chen, Boyan Li, Angel Oluwale-Rotimi, Tomy Pelletier, Jessica Song

Question 1

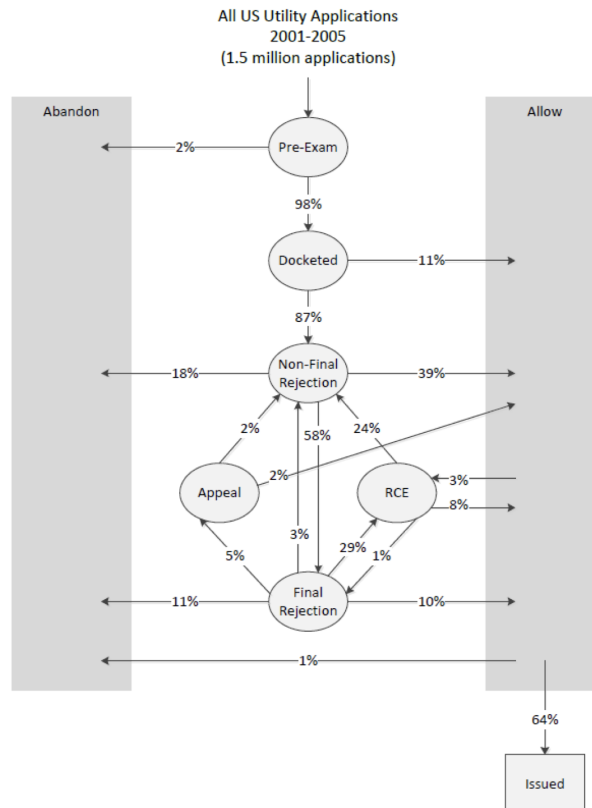
If examiners operate under the presumption that the majority of patent applications will inevitably undergo the non-final rejection process, this mindset can manifest into reality, affecting their efficiency and decision-making.

This behaviour could be further driven by the USPTO's examiner production measurement system. Within the framework of the biweekly production tracking, a non-final rejection is valued at 1.25 production counts, incentivizing examiners to lean towards this rather than the more time-consuming process of a thorough review for potential allowance - which may extend beyond the two-week period and not earn any counts in that period.

Figure 1: Representative Application Flow Diagram from USPTO Patent Prosecution and Examiner Performance Appraisal primer by Marco et al.

```
getwd()
```

```
## [1] "/Users/aoluwolerotimi/GitHubProjects/Talent-Analytics-2024"
```



Question 2

Propose an embedding-based measure to evaluate the presence of the mechanism you have identified.

The proposed embedding-based measure hinges on the hypothesis that examiners perpetuating this self-fulfilling prophecy may resort to standardized, boilerplate language in their decision documents. Such documents would exhibit a high degree of text similarity.

The embedding-based similarity measures for the non-final rejection, the allowance, and the final-rejection can be done at the examiner level for a more controlled comparison. These can serve as a measure of how detailed of a review was performed to produce the decision. Subsequently, an aggregate measure of those scores at the document type level can be created.

A statistically significant difference in the similarity of the non-final rejection vs each of the other document types could stand as evidence of the self-fulfilling prophecy

Question 3

A form of inequality that could arise from this is that examiners who are handling more unique or complex applications could earn lower productivity metrics because of more detailed reviews of new applications on their docket.

To help reduce this inequality, an LLM could be developed that assesses patent applications for uniqueness and complexity through two primary methodologies: text similarity analysis and topic modeling. For assessing uniqueness, the LLM would compare the text of a given application against a comprehensive corpus of prior applications within the same Technology Unit, identifying degrees of novelty. For complexity, the diversity of topics and topic distributions would be compared. If a meaningful elevation of either is identified, then a similar analysis takes place for examiner decision documents. The decision documents produced by examiners would be compared against a corpus of their previous decisions to evaluate the relative uniqueness of their responses. Similarly, the examiner's decision document would be screened for a meaningful relection of the topics presented in the patent application.

Incorporating these metrics into the existing quality component of examiner performance evaluations offers a mechanism to counterbalance the potential depression of productivity metrics. Examiners dedicating necessary attention to novel or intricate applications would receive recognition for their detailed review process and vice versa, aligning quality and productivity in performance assessments.

Properly defining the conditions for a meaningful elevation of uniqueness or complexity, which triggers the adjusted evaluation of the examiner's work product, is critical for ensuring that this does not end up penalizing efficiency in docket management. Examiners should neither be penalized for thoroughness nor incentivized towards overly elaborate reviews.