**Final Report: Social Sustainability in the Professional Journey**

**IT3389 Applied AI Project | Nanyang Polytechnic | Year 3**



| ***Team member’s name*** | Siew Wei Heng | Ethan Lee | Lim Hon Lin Gavin |
| --- | --- | --- | --- |
| ***Admin Number*** | 220342H | 220664H | 224729C |
| ***Roles played*** | Team Leader | Team Member | Team Member |

# Table of Contents

# 

[**Table of Contents 2**](#_js611enqj8rm)

[**Solution 3**](#_3726ue40ysxe)

[Salary Prediction (Wei Heng) 3](#_21mo4l9tr97p)

[Post Predictor (Ethan) 3](#_krxgp6suuyq)

[Education (Gavin) 3](#_nyos3ebmzu7x)

[Solution conclusion (Team) 3](#_n9m9kpktoaii)

[**Model & Accuracy 4**](#_aipbhrkv4581)

[Salary Prediction (Wei Heng) 4](#_8p517qf26f9g)

[Post Predictor (Ethan) 4](#_f6rp2f581hv9)

[Education (Gavin) 4](#_syara6l24ti8)

[Model & Accuracy conclusion (Team) 4](#_9t0ze58m4pvi)

[**Features 5**](#_9q6gx75f9btt)

[Salary Prediction (Wei Heng) 5](#_qioddolg7tdg)

[Post Predictor (Ethan) 5](#_tw8ltxhnffew)

[Education (Gavin) 5](#_bkti4dkvvwph)

[Feature conclusion (Team) 5](#_urm2m144052u)

[**UX 6**](#_b9vkmcej4ske)

[Salary Prediction (Wei Heng) 6](#_njyojx9ggo6g)

[Post Predictor (Ethan) 6](#_gg3r4qk2btmv)

[Education (Gavin) 6](#_avuqb28v1sxy)

[UX conclusion (Team) 6](#_emm8yeejydf2)

[**Reflection 7**](#_gnzkdzuzku82)

[Salary Prediction (Wei Heng) 7](#_s1m768yiika9)

[Post Predictor (Ethan) 7](#_szoq9vt3gcns)

[Education (Gavin) 7](#_tpkv69p5cfb8)

[Reflection conclusion (Team) 7](#_hh7esvilsdio)

# 

# Document Rules (this section will be removed on submission)

* Arial font
* Size 11
* No unnecessary bolding of headers
* No unnecessary adjustments of H1 and H2 headers
* Only permitted to use H3 and below headers to segment text.
* There are no page limits, thus we don’t need appendices and figures. For simplicity, place tables, figures, links directly in the appropriate text.

# Solution



Salary predictions, Post predictions, and Education Predictions will be the order in which we go through our solution, as well as our other sections.

## Salary Prediction (Wei Heng)

The salary predictor solves the problem of high time and resource cost associated with research into other countries for career opportunities.

This is done by allowing users free access to our [web application](https://work-advisor.vercel.app/salary) in order to save enormous amounts of time in research. Users are able to key in their custom profile, and have predictions for multiple countries and multiple locations in a short amount of time, accompanied with an insightful written report including statistics, percentages, recommended insights, and further questions, from a chatbot.

The predictions are trained on 57K [job listings](https://www.glassdoor.sg/Job/united-states-python-developer-jobs-SRCH_IL.0,13_IN1_KO14,30.htm) from a variety of countries and locations. Users can expect to compare the US, Singapore and India for a wide variety of economic positions. The model has an acceptable error rate of +/- 20K USD /year, given the highly specific nature of the input, and its performance compared to state of the art language models.

## Post Predictor (Ethan)

In essence, the post predictor solves the problem of the lack of access to personalised guidance for professionals and students alike.

I deployed a predictive model and GenAI with fastapi on Google Cloud Run, and these applications can be interfaced through the [post predictor portion of the website](https://work-advisor.vercel.app/post-prediction).

The predictive model is a Neural Network model trained on posts made on this [forum](https://www.thestudentroom.co.uk/). Essentially, you can input the category, title and post content that you want to publish, and the model will predict the popularity of the post after 1, 3 and 7 days, to cater for different levels of urgency for the post. (Some people require quick responses ASAP, while others don’t mind waiting).

The Generative AI is an additional feature that is explained in detail below, and essentially prevents invalid posts from being predicted and provides feedback on how to improve the current post.

The post predictor addresses the general problem of unequal opportunities by ensuring that users can gain equal access to relevant and personalised guidance. By predicting which posts will garner the most engagement, it increases the likelihood that users receive the feedback they need, allowing for more informed decisions in their educational and professional paths. This enhances social sustainability by reducing barriers to useful information and promoting fairer access to career and educational opportunities.

## Education (Gavin)

To sum it all up, the education page aims to solve the questions that hover around in many people's minds. To go in depth,

Our solution was created to help users spot imbalances and uncover patterns of inequality within educational systems across different countries. It’s designed to offer a deeper understanding of the challenges that various nations face when it comes to education. By looking at important factors like enrollment rates, completion rates, proficiency levels, and unemployment, the platform shines a light on areas where attention is needed most, giving those in charge the information they need to take meaningful action. When users key in numbers prior to the respective fields, it will return a Visualisation as a prediction output for users to have a better grasp graphically.

This project has been combined with the work of my teammates, bringing together our strengths to create a complete and functional platform. After a lot of hard work, the model has been successfully launched and is now live. The web app is easy to use, allowing users to explore the data and see how different countries are performing in education.

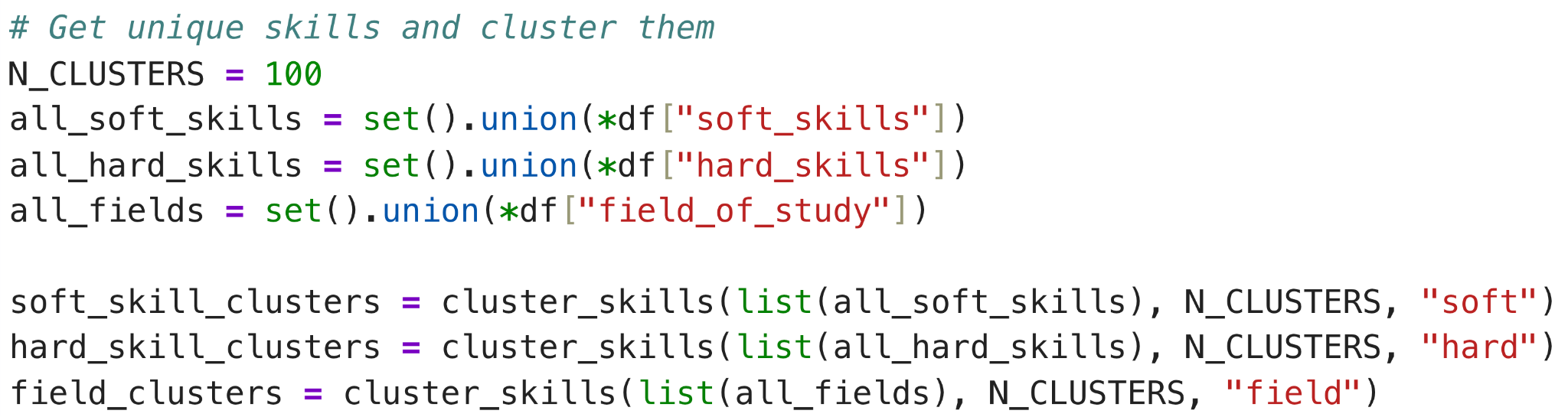
This isn’t just a tool — it’s about creating change. It highlights where education systems are falling short and provides users with the insights to push for improvements. With the model now live, the platform offers real-time information that helps raise awareness and encourages important conversations about how we can make education more fair and accessible. Through this project, we hope to inspire action and contribute to building a more inclusive and sustainable future for education around the world

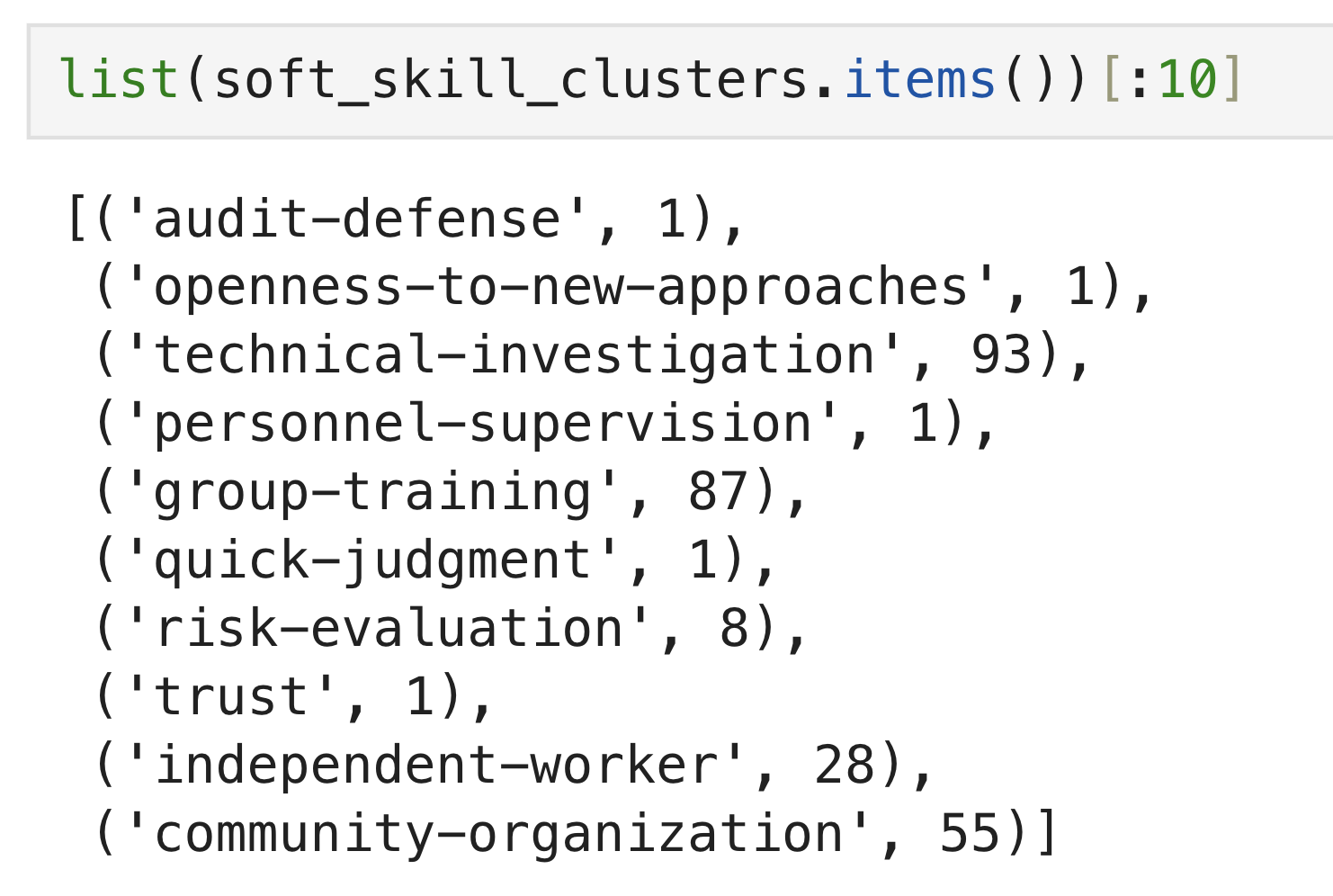
# 

# Model & Accuracy

## Salary Prediction (Wei Heng)

### Clustering



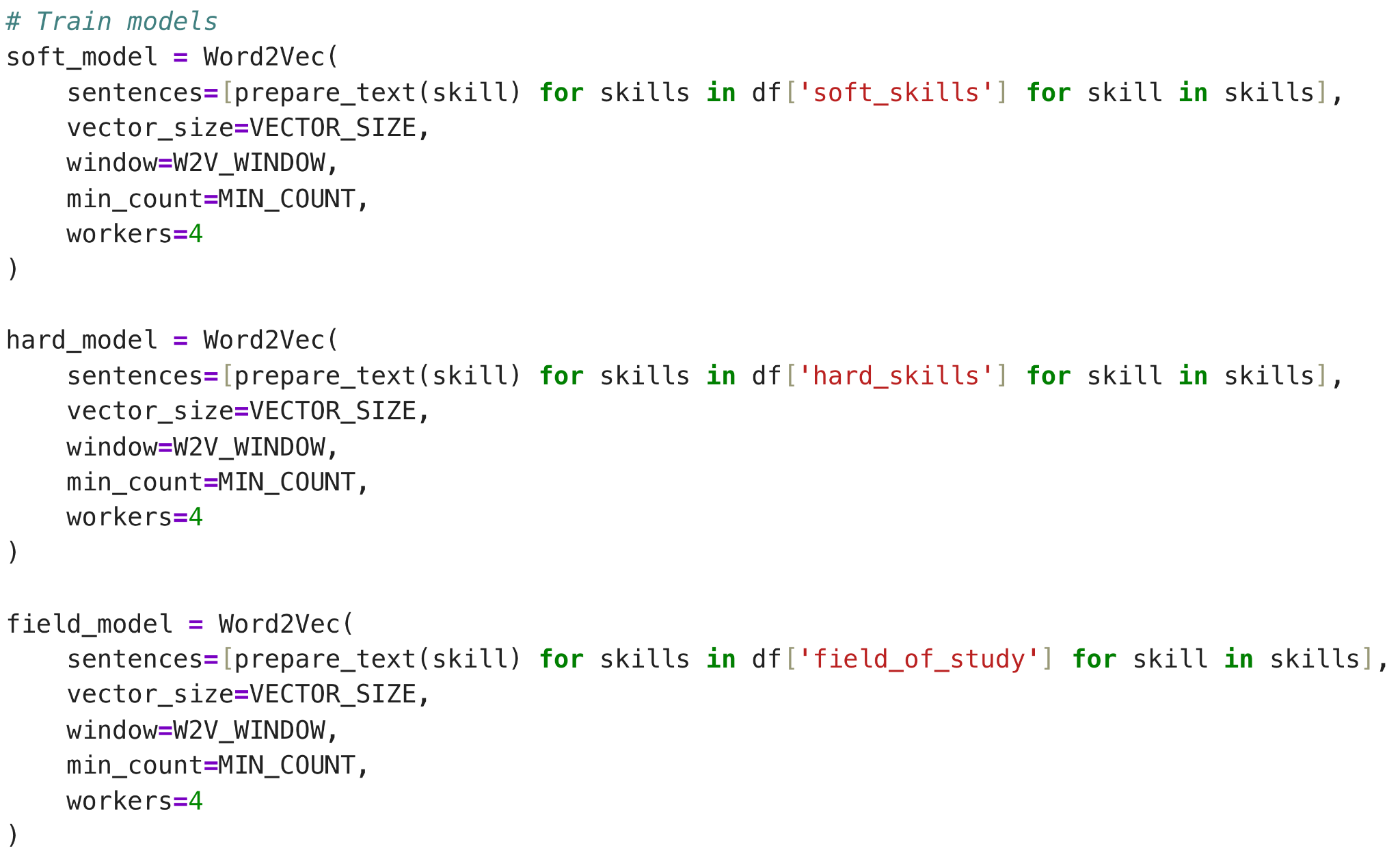


Clustering is done by assigning elements to a cluster based on how similar they are in terms of their TF-IDF scores.

The first element, ‘audit-defense’, represents the skill. the second element ‘1’ represents the cluster id.

These features are fed into a random forest model.

### Word2Vec



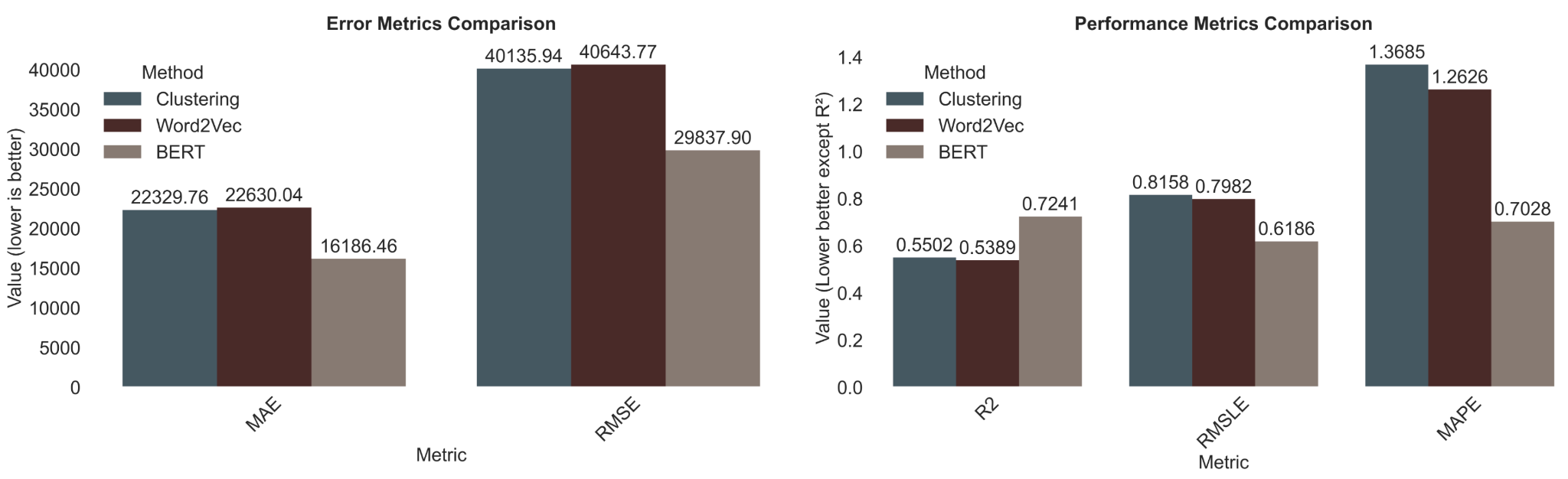
Word2Vec is done by training a Word2Vec model on soft and hard skills, and the field\_of\_study field, and feeding the features into a random forest model.

### DistilBERT feature extractor



The distilBERT feature extractor is simply the early layers of the DistilBERT model with a predictor at the end. The model is trained for only 5 epochs.

### Results

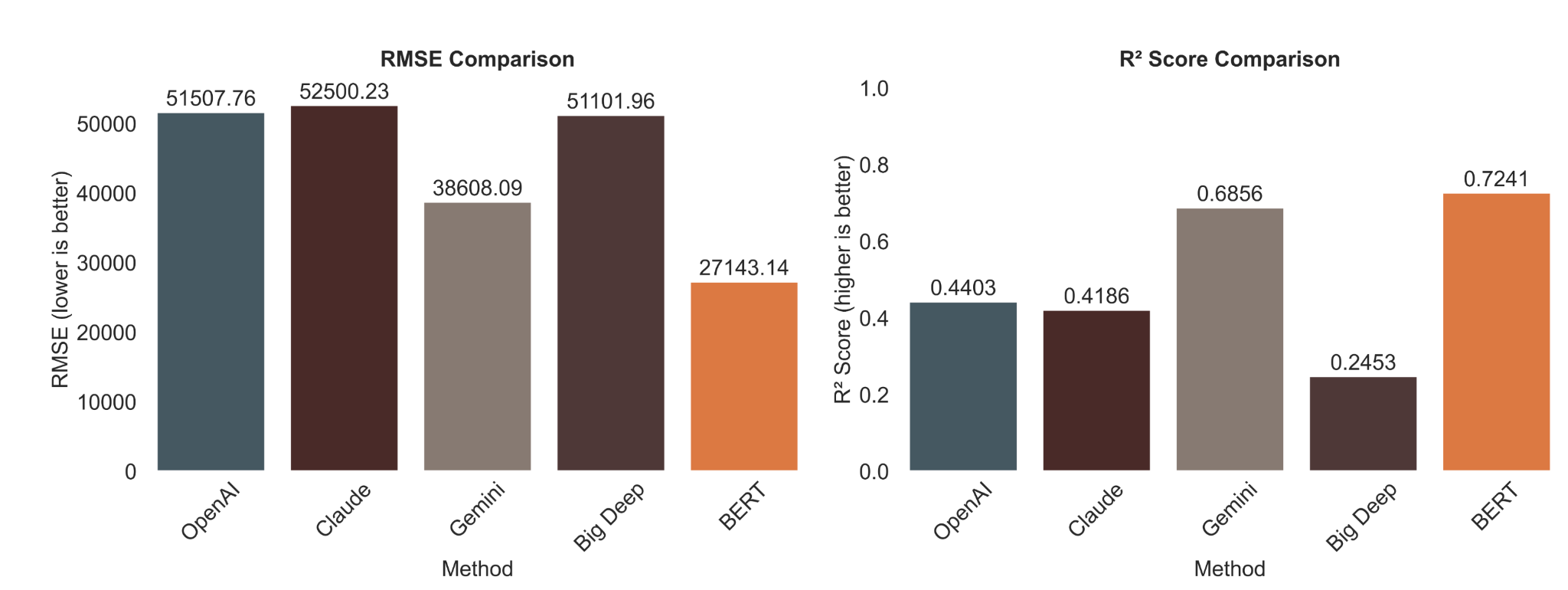


BERT has performed much better than the other non-deep learning models.

### Big deep

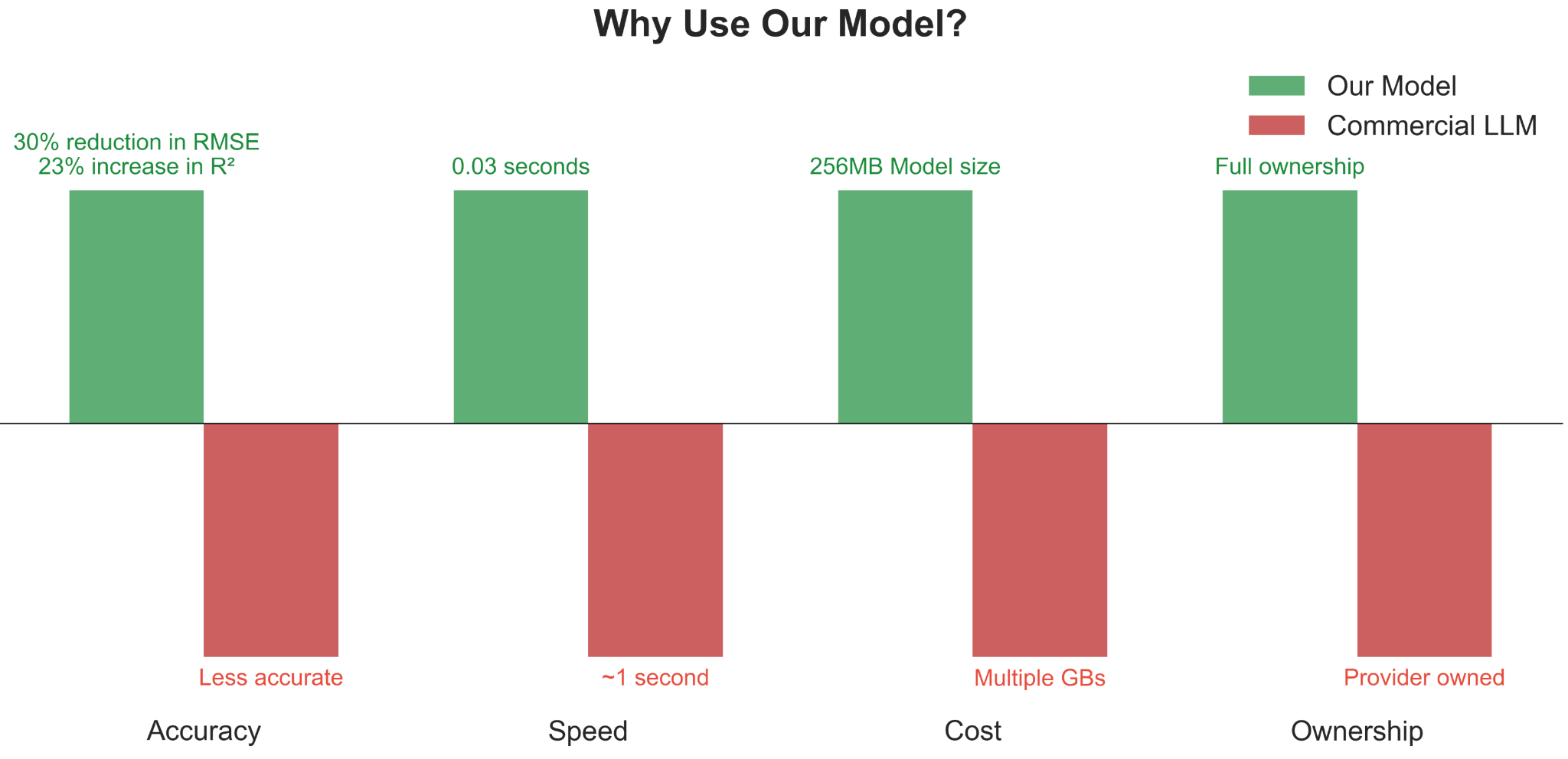
The big deep learning model is 469MB in size and features a very extensive model architecture.

### Results



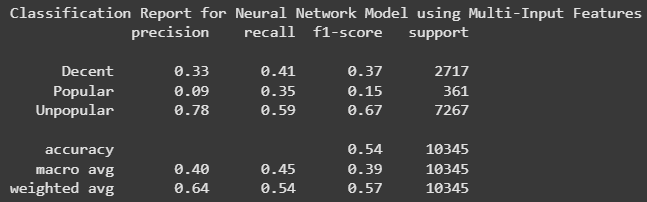
BERT has performed much better than the other deep learning models.

Surprisingly Gemini is very competitive.

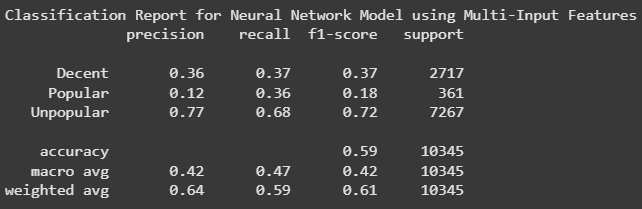


As such, we have deployed the BERT model to be used with our web application.

## Post Predictor (Ethan)



*First Neural Network (NN) Classification Report*



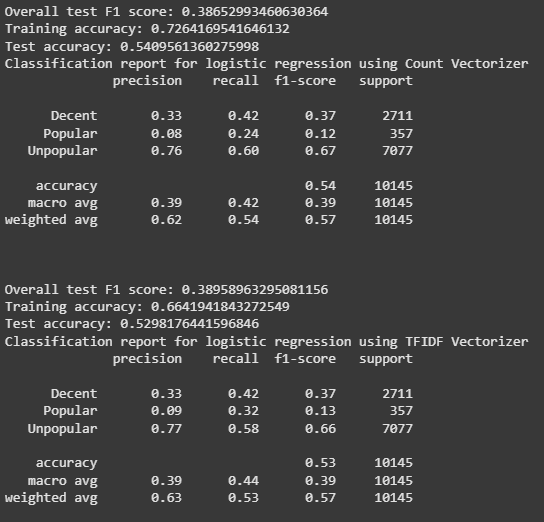
*Final Optimised NN Classification Report*

Above are the classification reports for both my neural network models.

All around, our priority is to increase the F1-score to balance identifying engaging posts while avoiding promoting low-quality ones. From the metrics, we can see that the new model maintains a recall above 0.33 for both Decent and Popular posts, ensuring that we continue surfacing engaging content.

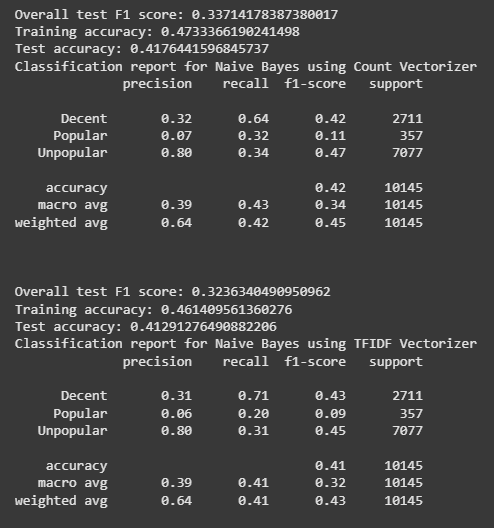
Additionally, the F1-score for Popular posts improved from 0.15 to 0.18, and overall accuracy increased from 0.54 to 0.59, demonstrating a more balanced performance.

The only issue is that precision for decent posts have decreased by .03, but the trade-off results in better recall for engaging posts, and an overall increase in accuracy. Ultimately, this model is stable, but its predictive capabilities for popular and decent posts seem to be limited. This could be improved with more data and optimising NN layers further.



*Logistic Regression Classification Report*

The drop in training accuracy and test accuracy suggests a significant amount of overfitting, but the results of the test are quite spectacular for both models in comparison to the neural network. We can see here that f1 scores for the “Decent” category are the same as the deep learning model, but they fall short for popular and unpopular categories. Although these models are comparable to the neural network model, the only improvement is recall for the “Decent” category for both models, and it is not justifiable to replace this model with the deep learning model. Overall, the f1 scores for both models are decent, but not as good as the neural network model.

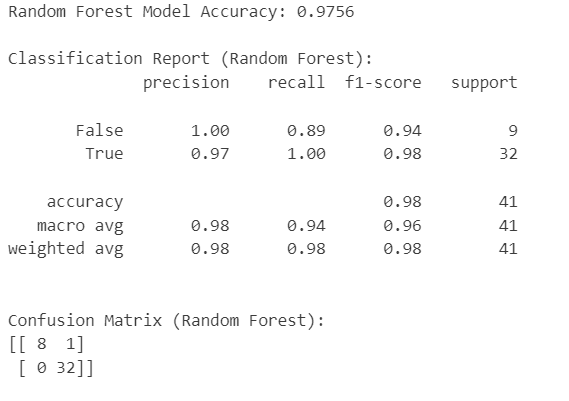


*Naive Bayes Classification Report*

There is clear overfitting for both Naive Bayes models. Outside of the high recall for the “Decent” category and high precision for unpopular posts, the other categories are horrendously predicted, and the overall f1 score is sub par at best. This definitely cannot replace the neural network model.

Ultimately, we will be sticking to the Deep Learning model and deploying it for the application, as it displays the best overall f1 score and balance for all classes.

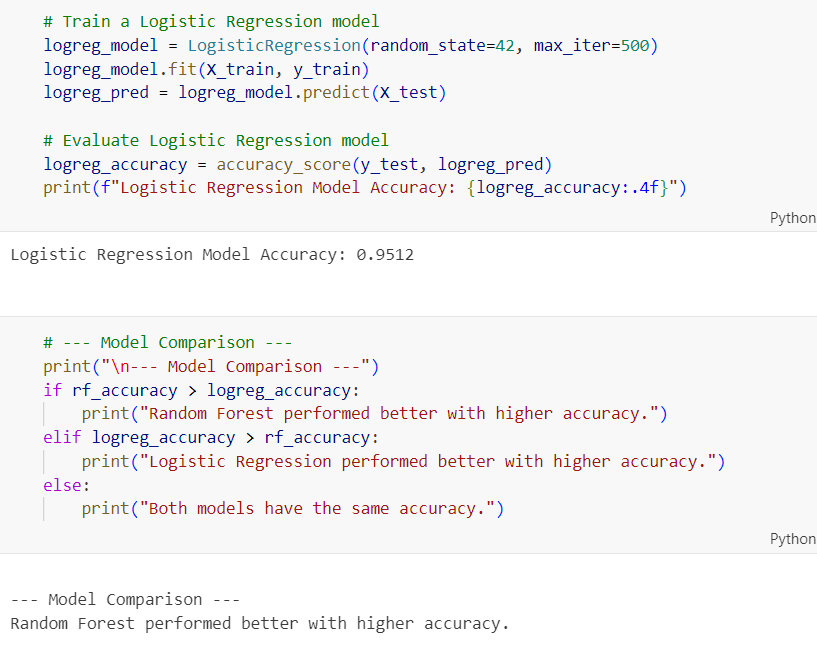
## Education (Gavin)



*Classification Report (Random Forest Model)*

## 

*Comparative Classification Report (Logistic Regression Model)*

**

*Comparative Classification Report (Final Result)*

1. (Classification Report): Random Forest Model allowed our model's overall performance on this dataset to be pretty optimistic. This is evident from the high accuracy (0.9756), perfect recall for the positive classes, and good F1-scores which also suggest a well-performing model albeit the small dataset size, with appropriately 202 rows and 31 columns.

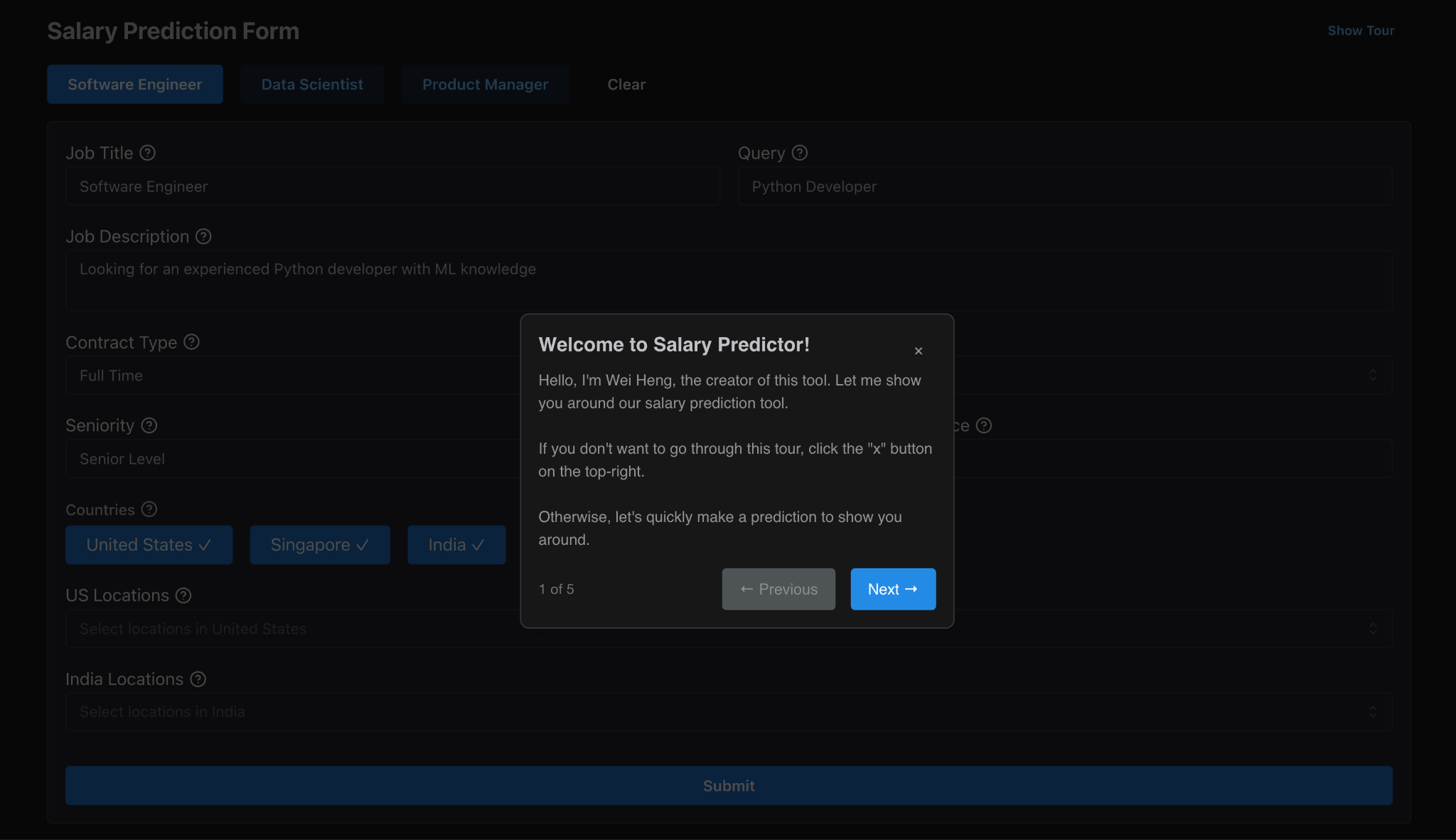
2. (Comparative Classification Report): Additionally, by comparing Random Forest Model with a Logistic Regression Model, it is shown that the Random Forest Model has a higher accuracy as compared to the Logistic Regression Model. This highlights better model performance, which explains why it is my pick.

# Features

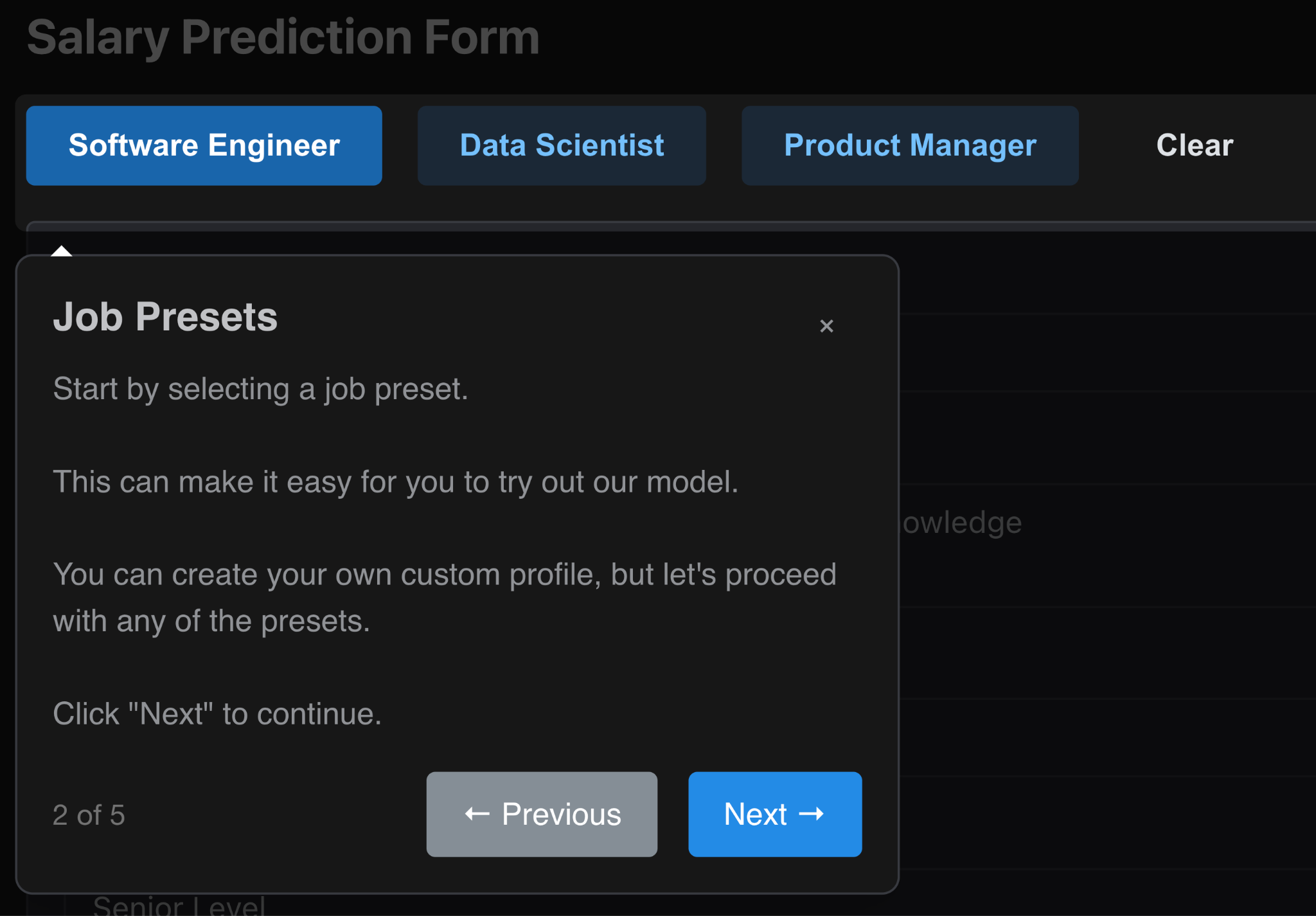
## Salary Prediction (Wei Heng)

The salary predictor has much more fields than a typical prediction scenario, which is why the form experience must be excellent.

**Feature 1: Product Tour**



The user is greeted with a tour that takes them through and explains what the tool can do, some context behind the model and prediction process, and guides them towards clicking the submit button.

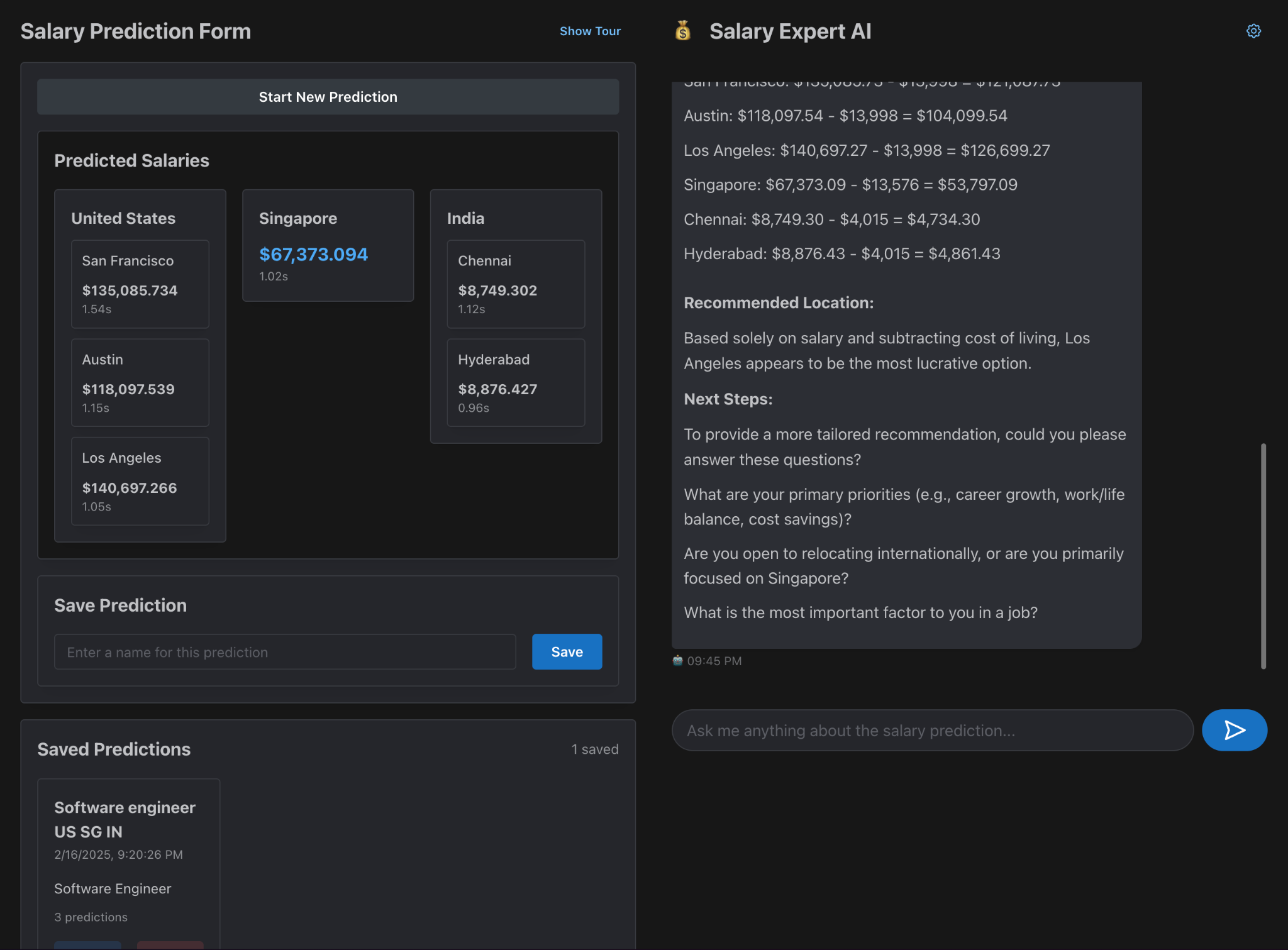


This tour allows new users to be familiar, and not to get overwhelmed by the customizability of the inputs.

**Feature 2: Iterative Predictions:**



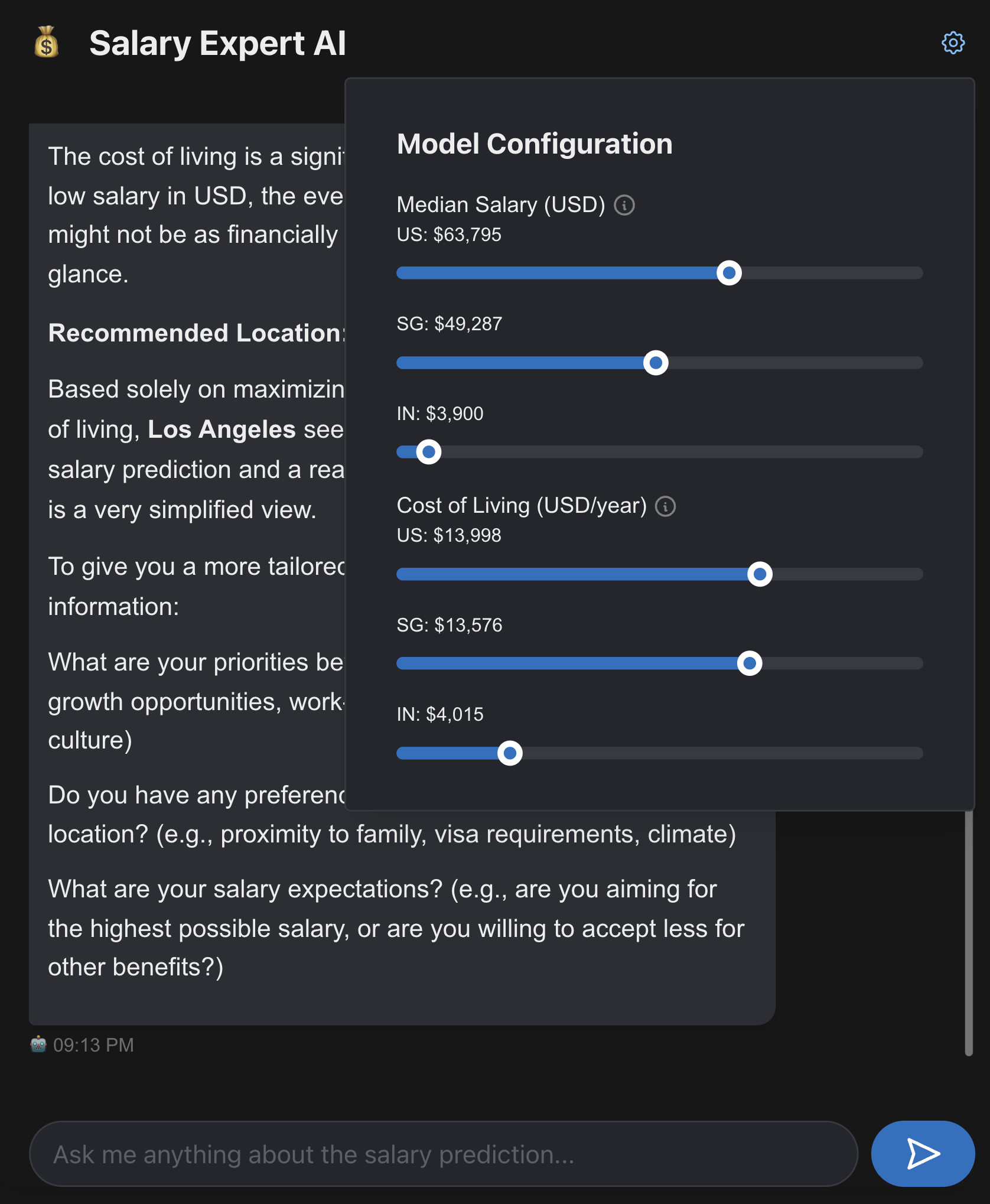
By selecting multiple countries and location, the model will do predictions for each location, allowing for the user and chatbot to compare opportunities.



Feature 3: Automated report analysis.

Based on the rich data provided by:

1. Your input details regarding the salary predictions
2. The predictions of the various countries, and…



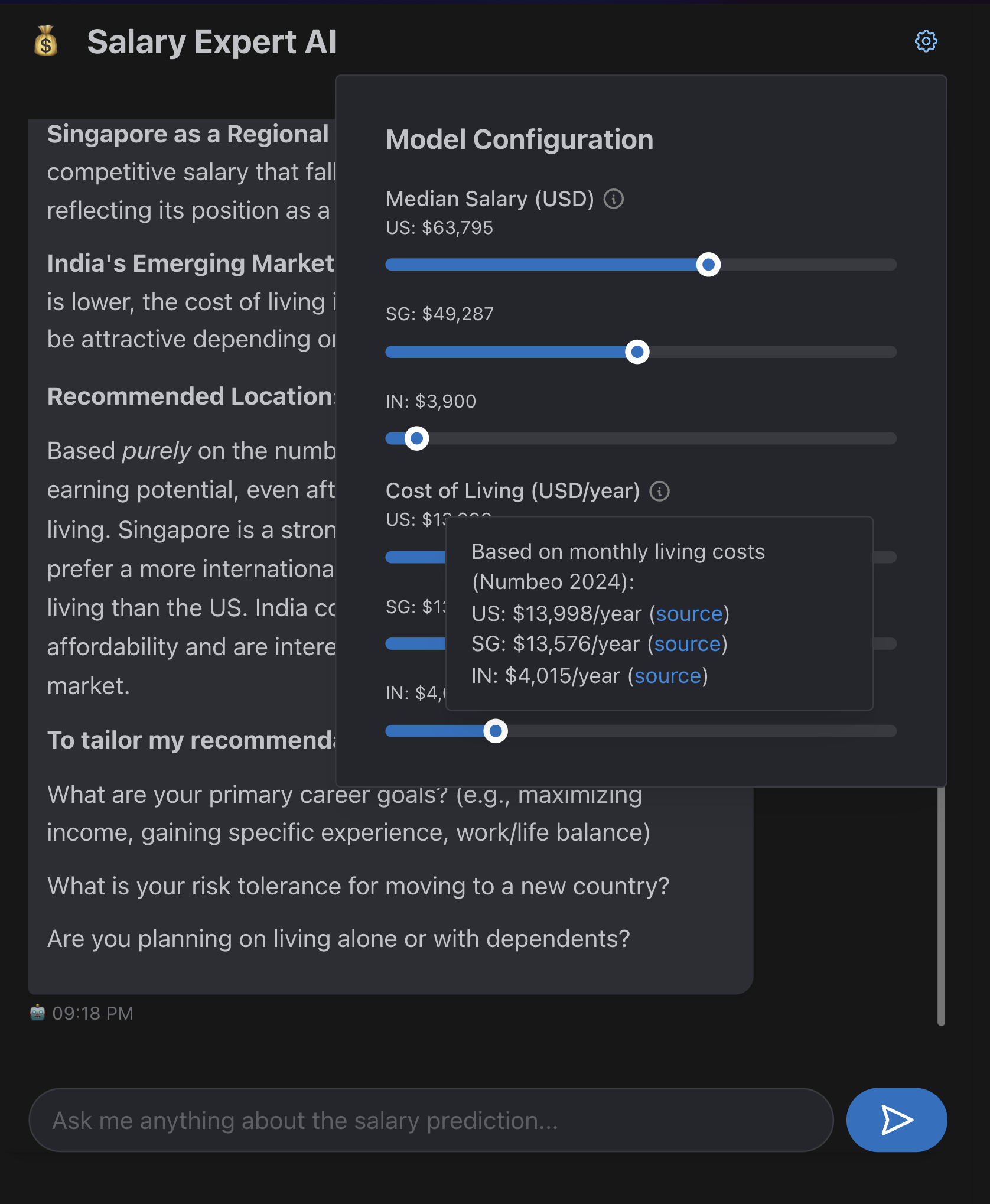
1. Metadata in regards to median salary and cost of living
2. General context like:
   1. The predictive model is trained with DistilBERT as the feature extractor, on glassdoor job posting data.
   2. The data is collected in Jan 2025.
   3. The model has an average error margin of approximately 20K USD/year.
   4. The user very likely currently resides in Singapore, though you're not 100% sure.
   5. etc…

The chatbot is well equipped to analyze and share findings together with the user.

Please visit our public github repository at <https://github.com/pclk/workAdvisor> in order to view interactive demos of the features.

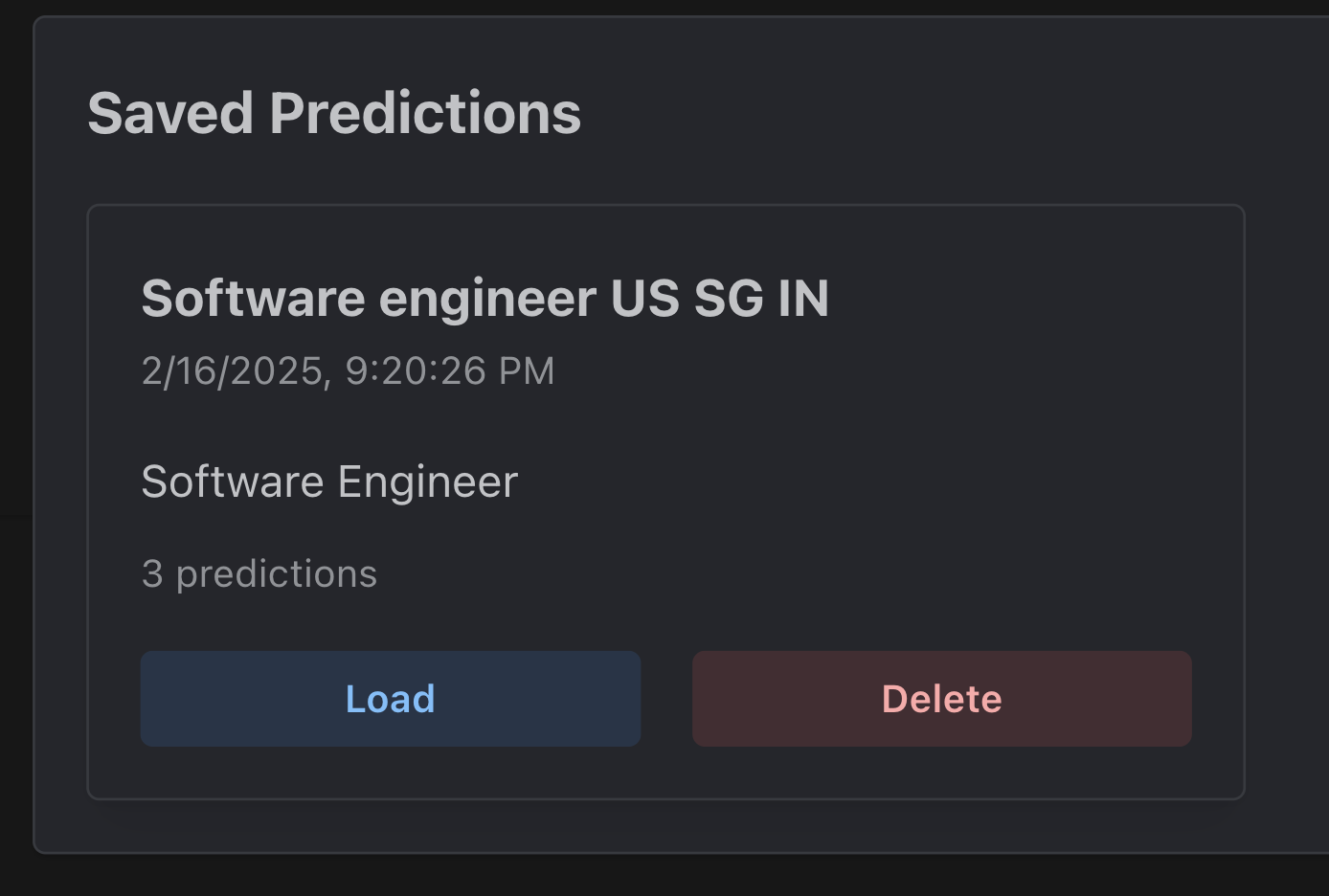
**Feature 4: Model Configuration**

As mentioned, we can adjust model configuration to allow for up-to-date information in regards to time-sensitive details like median salary and cost of living.

Sources are provided for the default values.

**Feature 5: Saving and Loading of predictions**





Users can saved their predictions as well as their input fields, and load them instantly whenever they want to.

**Feature 6: Chatbot**

The chatbot has conversation memory. That means after the chatbot provides you the analysis report, you can immediately load another prediction and chat with the chatbot to do an opportunity comparison. This can be extended to many predictions, limited by the chatbot’s context window.

## Post Predictor (Ethan)

We will go through 3 additional features for the post predictor, namely GenAI, Post Suggestion and Updating, as well as storage of analysis prediction history.



*Feature 1: GenAI*

The Generative AI component has the ability to filter information and determine if the post is not suitable for posting on the internet. It checks for personally identifiable information, nonsensical posts and posts that do not align with the category of the post. If any of these filters are detected, it mentions the issue above. Moreover, if these checks are flagged, the predictive model will not be called to save resources.

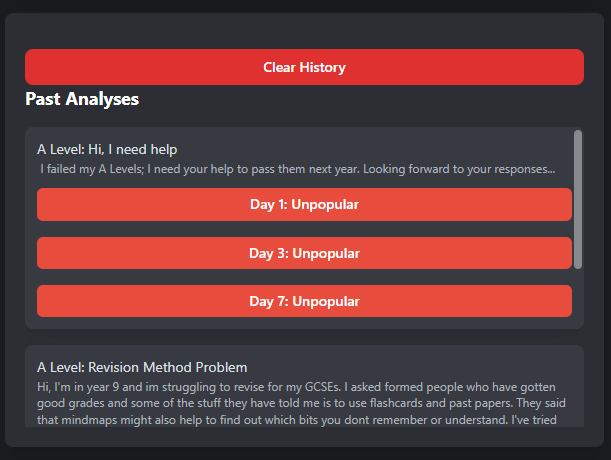
Additionally, the AI has the capability to generate recommendations on how to improve the post, as well as always provide the link to a page with posts of the similar category based on the current selected category. It also always provides outputs in point form (unless manipulated to do otherwise)



*Feature 2: Suggestion*

This feature allows the user to implement suggested changes from the AI immediately. After clicking on the ‘edit’ icon, the changes are placed into the post title and content. Moreover, the option to validate posts is temporarily disabled to prevent unnecessary calls to the model, since the changes are in an intermediary stage. If the user clicks on the tick icon, the changes are made. If the user clicks on the cross, the changes are reverted and the original post title and content are restored.

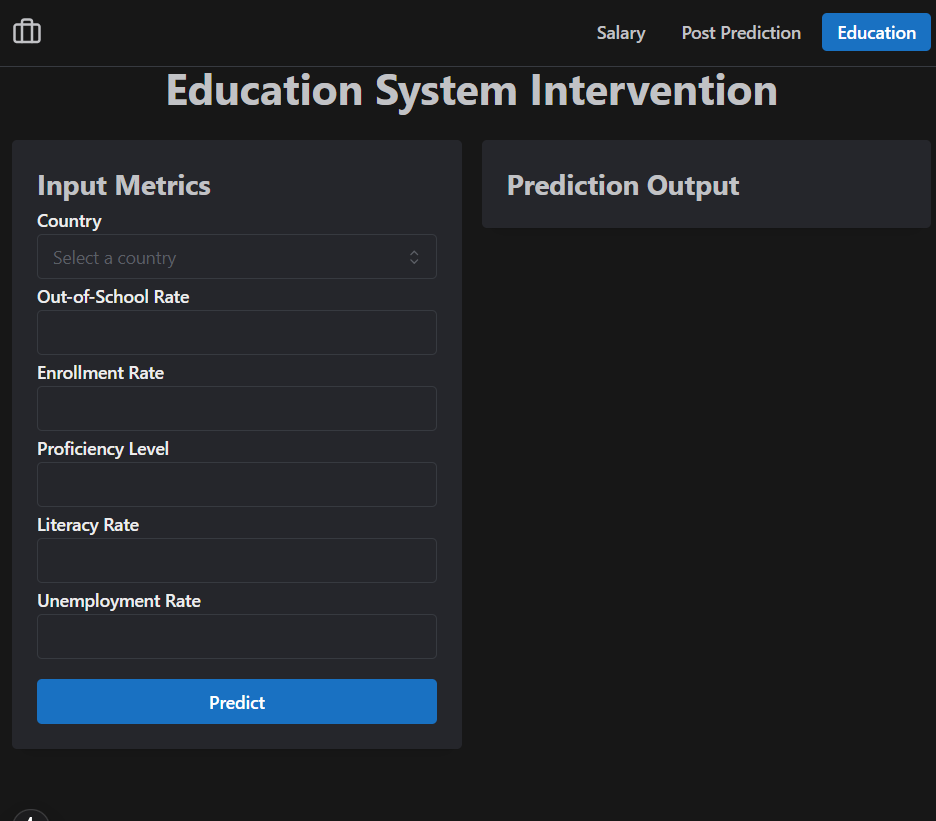
This feature removes the redundant manual and laborious task of having to incorporate each suggestion from the AI on your own, saving time.



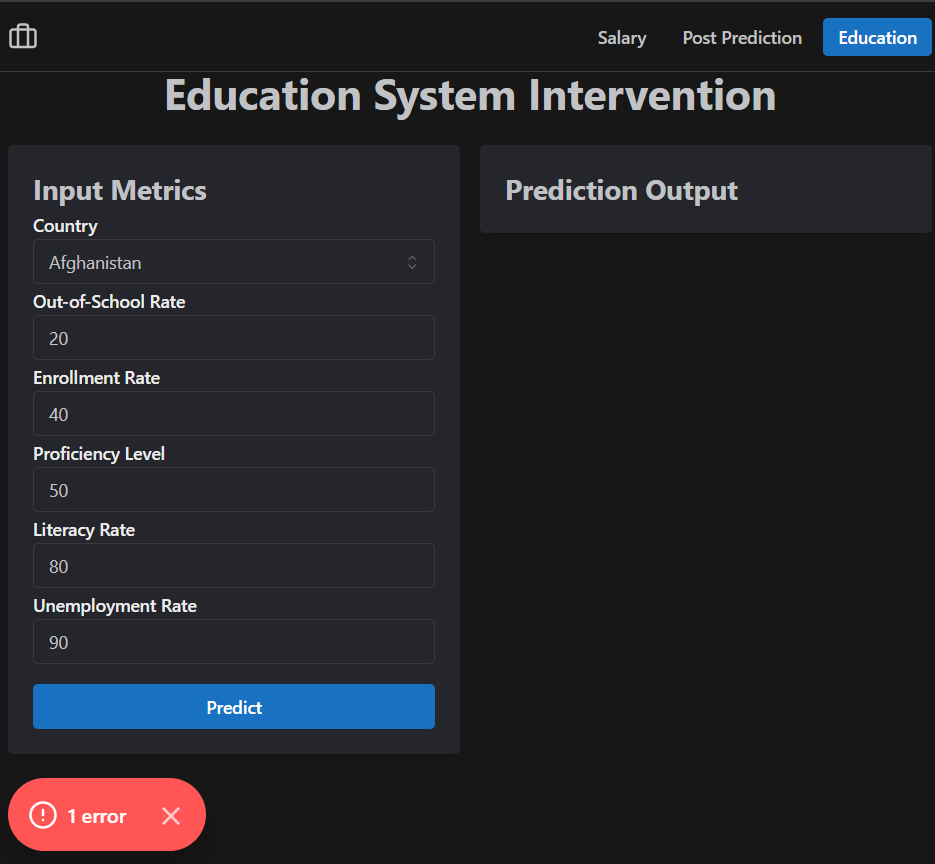
*Feature 3: Analysis History*

This feature allows the user to remember the posts they tried to predict. The history is stored in localstorage, and clearing history clears the history within localstorage (This is immediately reflected on the UI)

## Education (Gavin)



Over here, our objective is to allow users to enter inputs across the different countries in the world through a drop down alongside different metrics such as “Out-of-School Rate”, “Enrollment Rate”, “Proficiency Level”, “Literacy Rate” and “Unemployment Rate”. Upon filling in, our system will be able to return a Prediction Output which comes as a visualisation (Bar Graphs, etc - for the best display.) to identify trends - disparities in the educational system globally. This will allow users to have a bigger picture and grasp of the inequality in terms of education opportunities among countries. (Working on it.) After the visualisation is depicted, there will be an analysis panel which gives information & key details prior to the findings that they would want to gather and find out.



When Users input values that are out of the range accepted, it will prompt an “1 error” message at the bottom to indicate that Users are unable to key such values as it is inaccurate and wrong. There will be no result and no visualisations to be shown.

# 

# UX

## Salary Prediction (Wei Heng)

**UX principle 1: Aesthetic-Usability Effect (**[**Source**](https://lawsofux.com/aesthetic-usability-effect/)**)**

Users often perceive aesthetically pleasing design as design that’s more usable. That is why the salary prediction form, results and interface are designed with proper colours and spacing, allowing the user to feel at ease at every step of the way.

**UX principle 2: Chunking (**[**Source**](https://lawsofux.com/chunking/)**)**

A process by which individual pieces of an information set are broken down and then grouped together in a meaningful whole. This is effectively accomplished through the introduction of product tours into the tool. By introducing the tool chunk by chunk, users are less overwhelmed, and can get started using the tool sooner.

**UX principle 3: Cognitive Load & Choice Overload (**[**Source**](https://lawsofux.com/cognitive-load/)**) & (**[**Source**](https://lawsofux.com/choice-overload/)**)**

The amount of mental resources needed to understand and interact with an interface.

Earlier into the project, I have considered the possibility of introducing the chatbot as early as the form filling process to assist with form filling. However, I’ve decided not to implement it to respect the Cognitive Load tool. I believe that by introducing the chatbot too early, the user will be overwhelmed by the choice of deciding to fill the form by themselves or by chatting with the chatbot. They may also face significant cognitive load trying to understand how the Chatbot can help them in this tool.

Therefore, I have simplified the form filling process with a product tour and presets to choose from.

**UX principle 4: Doherty Threshold (**[**Source**](https://lawsofux.com/doherty-threshold/)**)**

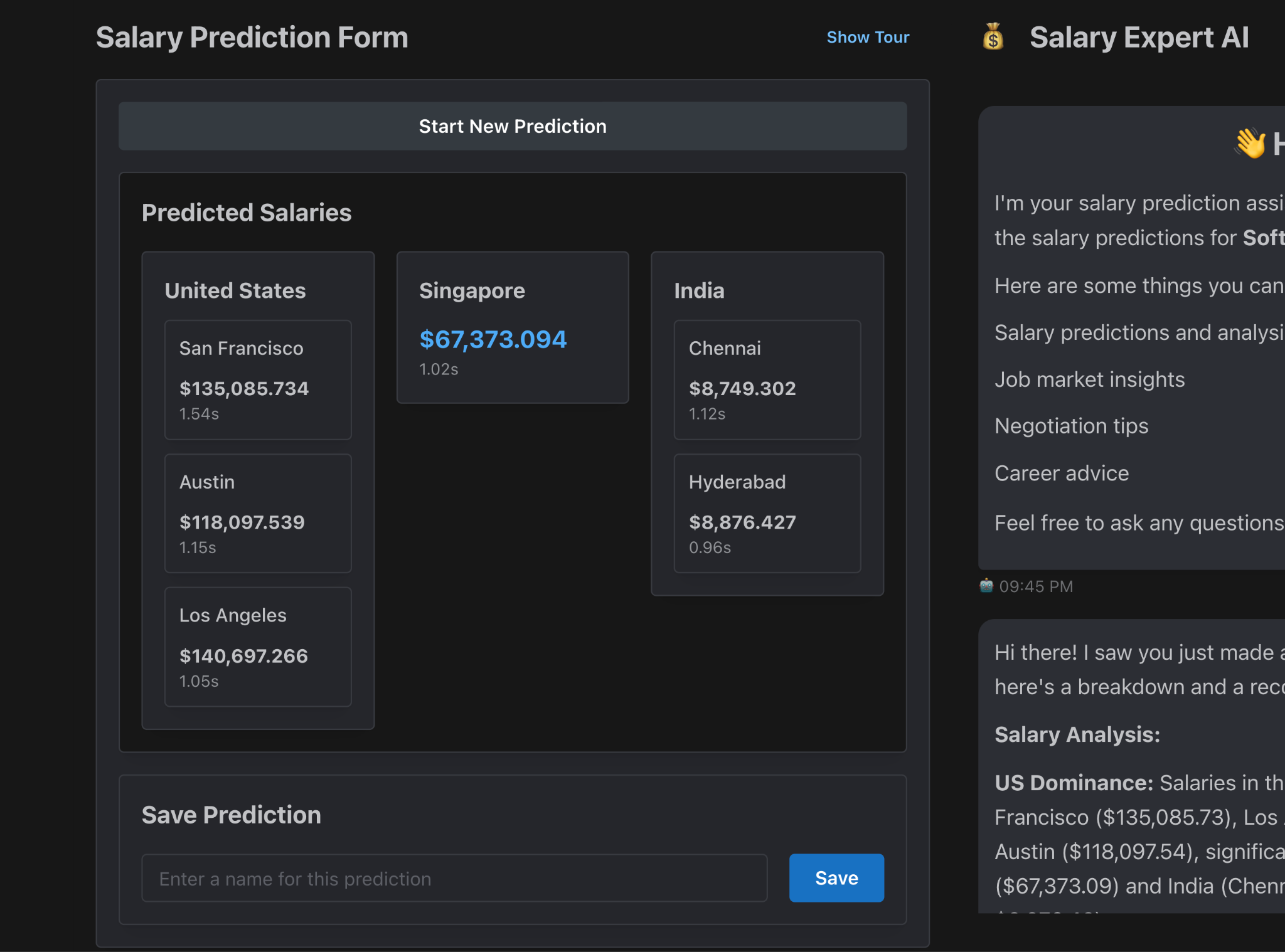
Productivity soars when a computer and its users interact at a pace (<400ms) that ensures that neither has to wait on the other.

The website and form has been designed with interactions and animations that make it extremely interactive and quick with the user. Some examples are a timer that quickly counts to a second, streaming predictions and chatbot responses, as well as detailed optimization of the interface reactive speed. This is achieved through advanced front-end technologies like state management, server-side rendering, site caching, and pre-fetching.

Therefore, the user wouldn’t have to feel stuck waiting on the tool, even though a powerful deep learning model is running in the background.

**UX principle 5: Law of Common Region (**[**Source**](https://lawsofux.com/law-of-common-region/)**)**

Elements tend to be perceived into groups if they are sharing an area with a clearly defined boundary. This is well demonstrated within the prediction display:



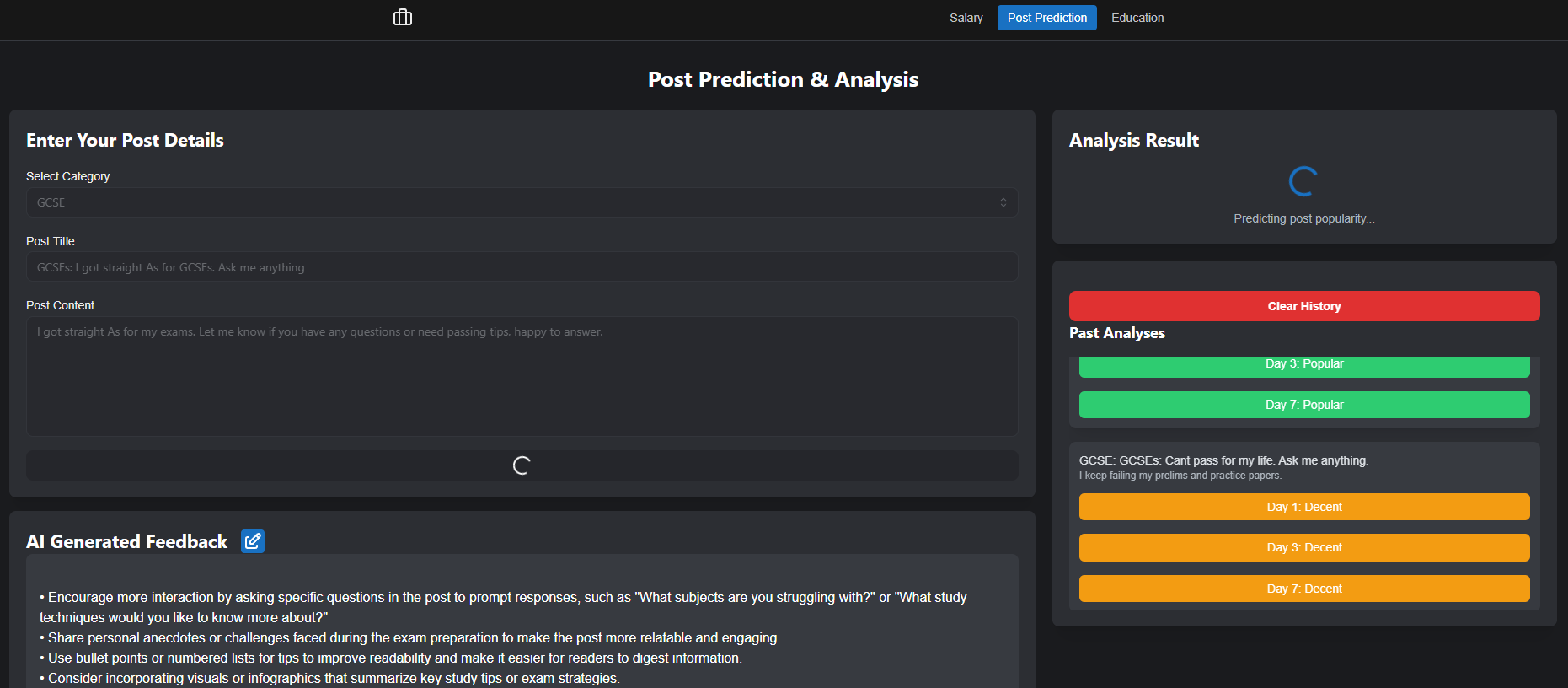
Information about a location are grouped together, which are grouped as location cards, grouped into countries, grouped into the predicted salaries portion.

## Post Predictor (Ethan)



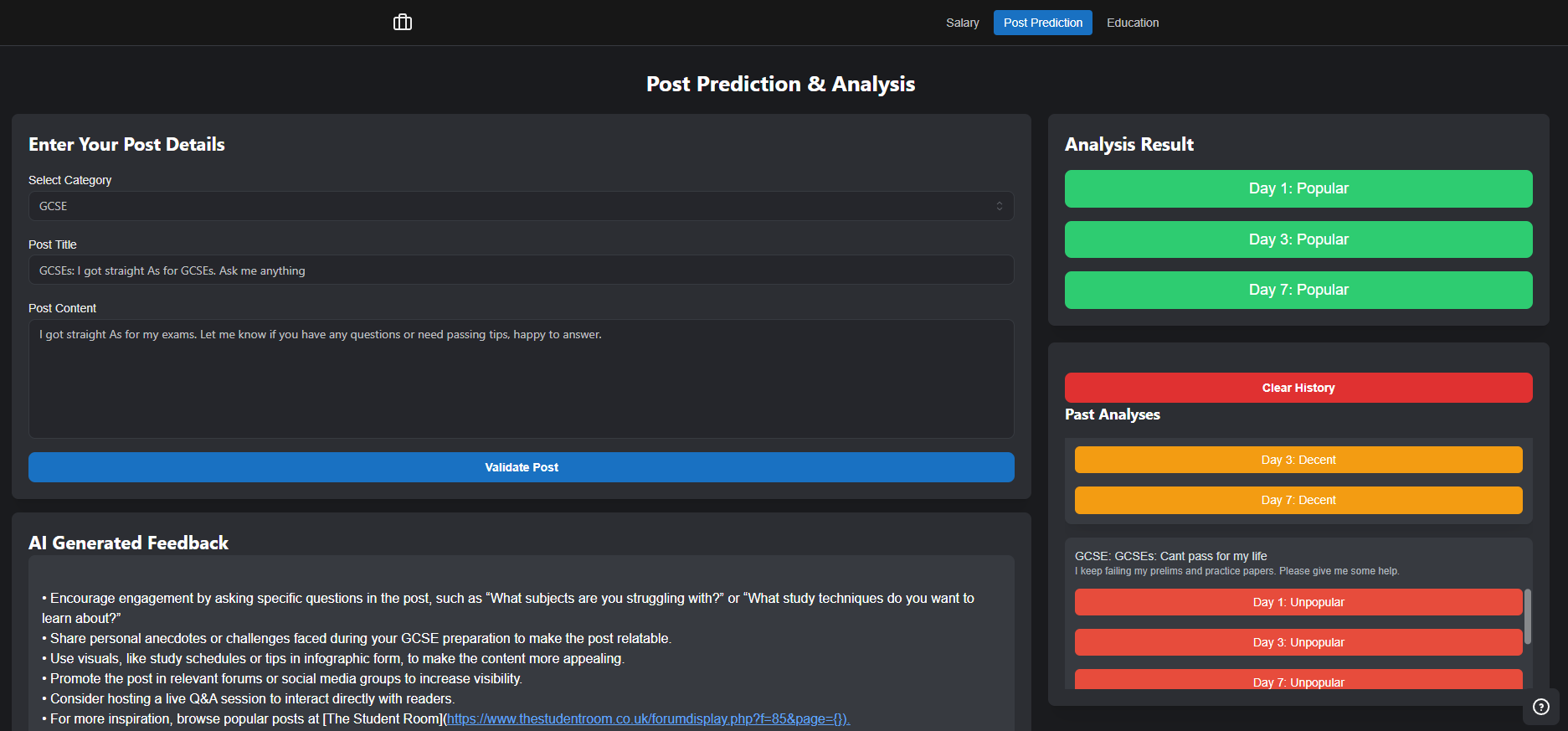
*Initial Page Tour*

The user experience is intuitive, featuring a guidance tour right off the get go. The tour highlights the relevant parts of the page and explains to the user what each part does. This occurs for post crafting, AI Generated Feedback, Suggestion Implementation, Analysis results, and Analysis History. Additionally, the user can click on the question mark highlighted in red to access the tour again.



*Page Loading State*

This prevents the user from re-submitting their post, and actively notifies them that their data is being processed. There is also clear feedback for prediction results, with colour coding (green for popular, yellow for decent, red for unpopular, helping users easily and intuitively understand feedback. Additionally, there are clear error messages using mantine’s notification component, alerting users if an issue is detected, such as a post containing PII (Shown above in [Features](#_tw8ltxhnffew)).



*Final UI (Zoomed Out)*

After going through the tour, the user will have clarity into what the page features and how to predict the popularity of their post. The figure above showcases an example of the final UI after the user has inputted the post they would like to predict.

## Education (Gavin)

To ensure a seamless and enjoyable experience, the design of the app focuses on simplicity and clarity. A clean, minimal layout allows users to quickly navigate through the features without feeling overwhelmed. Every element, from dropdowns to buttons, is clearly labeled, making it obvious what actions are being performed at all times. This intuitive design helps users focus on what matters most—the results.

Accessibility is at the core of the design. The interface is built to be fully navigable via keyboard, benefiting users with different abilities, while also speeding up the process for everyone. High contrast between text and background enhances readability, ensuring that the app is user-friendly for all. The responsive design ensures the platform works smoothly across devices, whether on a desktop, tablet, or mobile phone, making it accessible to users anywhere.

The interface is also visually engaging, with a consistent color scheme that guides users' attention to the most important sections. Key information is given priority, with size, color, and contrast making it stand out. This careful balance ensures that users can easily find and interact with critical elements, like the prediction outputs, without distraction.

Ease of use is paramount, and dropdown menus for input data streamline the process, allowing users to quickly make selections without wasting time typing. Immediate feedback on actions reassures users that their inputs are being processed, while the app's quick response time ensures results are delivered without unnecessary delay. All interactions are designed to be smooth and efficient, allowing users to get to the information they need with minimal effort.

The layout follows a logical structure that feels natural, with input fields at the top, followed by visualizations of the predictions and results. This consistent arrangement builds familiarity, so users always know what to expect. The app’s attractive design, with thoughtfully chosen fonts, icons, and charts, adds to its appeal, making the experience not only functional but enjoyable.

Errors are minimized with intuitive features like dropdowns, which reduce the risk of incorrect data entry. If mistakes do happen, the app provides helpful, clear messages that guide users toward correcting them without frustration. Furthermore, users are given the flexibility to personalize their experience, whether by adjusting settings or choosing their preferred themes, creating an environment that feels tailored to their needs.

# Reflection

## Salary Prediction (Wei Heng)

### **Lessons Learned**

The most difficult process of my portion in my opinion is the data collection process. While many of my classmates have collected well cleaned, prepared data from the internet repositories, and have many references to notebooks of the same datasets, I had to collect fresh data from scratch, because while there are synthetic salary prediction data, I would like to not only train my models on actual real life data, but also to give myself a challenge. Scraping internet job listings that are very well protected is not easy at all, and has consumed much of my time and energy. This is made even more difficult considering I had to collect 57k+ job listings to feel confident that deep learning networks can have enough data to train on.

Dealing with dirty, unpredictable data was very scary to me, because a wrong move could set my project back for multiple weeks. This was an enormous risk I took, one that I felt was worth it because I can proudly share that my model has true utility and is not simply a school experiment.

I have learnt a lot from taking this risky approach. I learnt to think out of the box to engineer features, which features to collect, and what modelling technique to utilize for my unique dataset. It is well known that students who decide on collecting their own data to train a deep learning model are on a set path to failure, but I’m truly happy to see that I’ve made it and outperform state-of-the-art language models trained on data millions larger than what I’ve collected.

### **Suggestions to improve solution**

I would start by collecting data outside of Glassdoor as well. Initially, I have identified multiple job listing sites to scrape, some of which are sites dedicated for one specific country, which could unlock valuable and variable data for my model to learn on. However, given the difficulty of scraping job listing sites, having to deal with cloudflare’s anti-scraping measures, and the limited time I have, I’ve decided to stick with Glassdoor.

Then, I would consider even more feature engineering techniques. Namely, which fields should be ordered first in BERT’s prompt? How to consolidate the text? This is to further reduce the RMSE of 20K+.

Finally, I would look at further optimizing the web application. If given the time, I would further improve on the user experience by further chunking the form into separate parts, and adding charts to display and compare the predictions, with automated calculations of median salary and cost of living, and possibly tax rates.

## Post Predictor (Ethan)

### **Lessons Learned**

I came across issues with scraping many times, as my Google Colaboratory notebook kept disconnecting after forgetting to interact with the tab or leaving models running overnight. This was so frustrating that I decided to actively reduce the number of rows I had to collect for scraping. This eventually would lead to lowered accuracy scores, which is regrettable now, but it teaches delayed gratification. I now realise that I could use an autoclicker of sorts to ensure interactivity with the tab, which I will implement if needed for future use. However, this goes to show how impatience and instant dopamine causes me to make suboptimal decisions, which I will refrain from making in the future.

I also unfortunately accidentally missed the deadline for the milestone report. This certainly broke my heart, but alas, nothing can be done about it. This taught me not to take things for granted, as I initially had the misconception that the milestone report was the actual presentation. However, I was not aware that we had to submit our models and application, which I regret, but am grateful for, as I will make sure I never suffer the same tragedy again.

Finally, in terms of technical challenges, it took a while to get used to performing pull requests on Github and learning the format of the React framework, but I got a hold of it and produced a satisfactory application. There were many setbacks which were self-inflicted, but they helped me grow and improve.

This project introduced many new concepts which I barely had practice with like AI ethics and Deep Learning, but ultimately, the experience gained, as well as the lessons learned, are invaluable. I can confidently say that if I were to tackle this project again, I would do more efficient work.

### **Suggestions to improve solution**

In terms of scraping, I believe that I should have scraped posts for multiple days, comparing the engagement levels for posts and adding extra columns for engagement popularity after x days. This would have been a more accurate measure of engagement prediction.

For EDA, I also regret not considering feature extraction for my dataset. I could have added additional features like a boolean for if the post is grammatically sound, or if the post had inappropriate language, which could have helped with the model f1 scores.

Additionally, as mentioned earlier, I would have scraped more data if I had more patience, as I had to restart my jupyter notebook many times. These would certainly have improved the existing solution.

In terms of solution modification, I would likely also have refrained from adding the “Decent” engagement category, as it has little difference from popular posts. Both simulate enough conversation to obtain personalised guidance, and combining both would have made the model much more reliable in its predictions.

## Education (Gavin)

Before I continue on with my reflection, I would love to dive back to how this project was truly more than just a technical challenge— it was a journey of personal and professional growth. By identifying gaps in education through data-driven insights, we are contributing to a cause that aligns with social sustainability, equity, and the belief that education should be accessible to all.

Now, here are some of the reflection that I have gathered throughout this process:

### **1. Dataset Challenges (Resilience & Problem-Solving)**

Finding the right dataset was one of the most difficult and frustrating parts of this project. I had to search for, evaluate, and refine multiple datasets before finding one that met the project's needs. It was stressful and time-consuming, requiring patience, critical thinking, and problem-solving. Despite the setbacks, I refused to give up, knowing that a strong dataset was the foundation of an effective model. This experience taught me resilience and how to adapt when things don’t go as planned.

### **2. Teamwork & Collaboration (The Power of Teamwork)**

This project would not have been possible without teamwork. Our leader, Wei Heng, and teammate, Ethan, consistently provided updates and support whenever I faced challenges. Their willingness to collaborate, offer solutions, and maintain clear communication made everything run smoothly. I learned the importance of team synergy—having a supportive team made even the toughest obstacles manageable. We divided tasks efficiently, played to our strengths, and encouraged each other to push forward.

### **3. Time Management**

Juggling multiple aspects of this project—data preparation, model training, and deployment—taught me the importance of prioritization and structured workflow. Some tasks, like data exploration, took longer than expected, but planning ahead and tracking progress helped keep things on schedule. This experience strengthened my ability to manage deadlines under pressure and stay focused on the bigger picture.

### 

### **4. Cloud Deployment Challenges (Technical Growth)**

Deploying the model to Google Cloud came with its fair share of hurdles, including authentication issues, versioning conflicts, and API optimizations. It required a mix of troubleshooting skills, technical research, and persistence. Overcoming these challenges reinforced my ability to debug complex systems, work with cloud-based services, and ensure smooth deployment.

### **5. GitHub & Version Control (Technical Growth)**

Managing version control was another steep learning curve, with merge conflicts and dependency mismatches causing unexpected issues. I had to quickly adapt, learning how to use structured branching as well as proper commit strategies. This process strengthened my technical workflow and reinforced best practices in software collaboration.

## Reflection conclusion (Team)

This project utilized Artificial Intelligence to address areas of inequality with a focus on sustainability. The team developed AI-driven tools for salary prediction, post engagement optimization, and the analysis of educational disparities, aiming to empower professionals and students. The project resulted in tangible technical outcomes and provided valuable learning experiences for the team members.

The project culminated in a web application that demonstrates practical utility and a coherent design. Significant features include a chatbot interface for the Salary AI model, GenAI suggestions and feedback for the post predictor, and a user-focused application design. The team achieved a high level of accuracy in salary prediction through data engineering and model development.

Each team member played a vital role and overcome specific challenges. Wei Heng focused on salary prediction, improving model accuracy through research and data engineering. Ethan addressed technical issues and adapted to new technologies, gaining experience with front-end development and collaborative workflows. Gavin worked with complex datasets and cloud deployment, developing solutions for educational disparity analysis and enhancing his independence in model development and deployment.

This project represents a contribution to the application of AI for positive social outcomes and sustainability. The development of practical tools addressing real-world inequalities, combined with the technical and professional development of the team, highlights the project's value and outcomes. The skills and experience gained are expected to be beneficial for future work in AI and its application to societal issues.