



Summer 2024 Project: Content Evaluator

Michelle Han

Overview

- Current problems
- ATS foundation
- Grammarly
- Project steps
- Logistics
- Obstacles

Current Issue

- Inconsistency with information on Developer Portal
- Many components to oversee
- Unclear guidelines
- More future content to be added
- Outdated current information
 - Example: Coleman AI



How to Import Data in
Coleman AI

Application Tracking System (ATS)



Review our suggestions to see what you can fix.

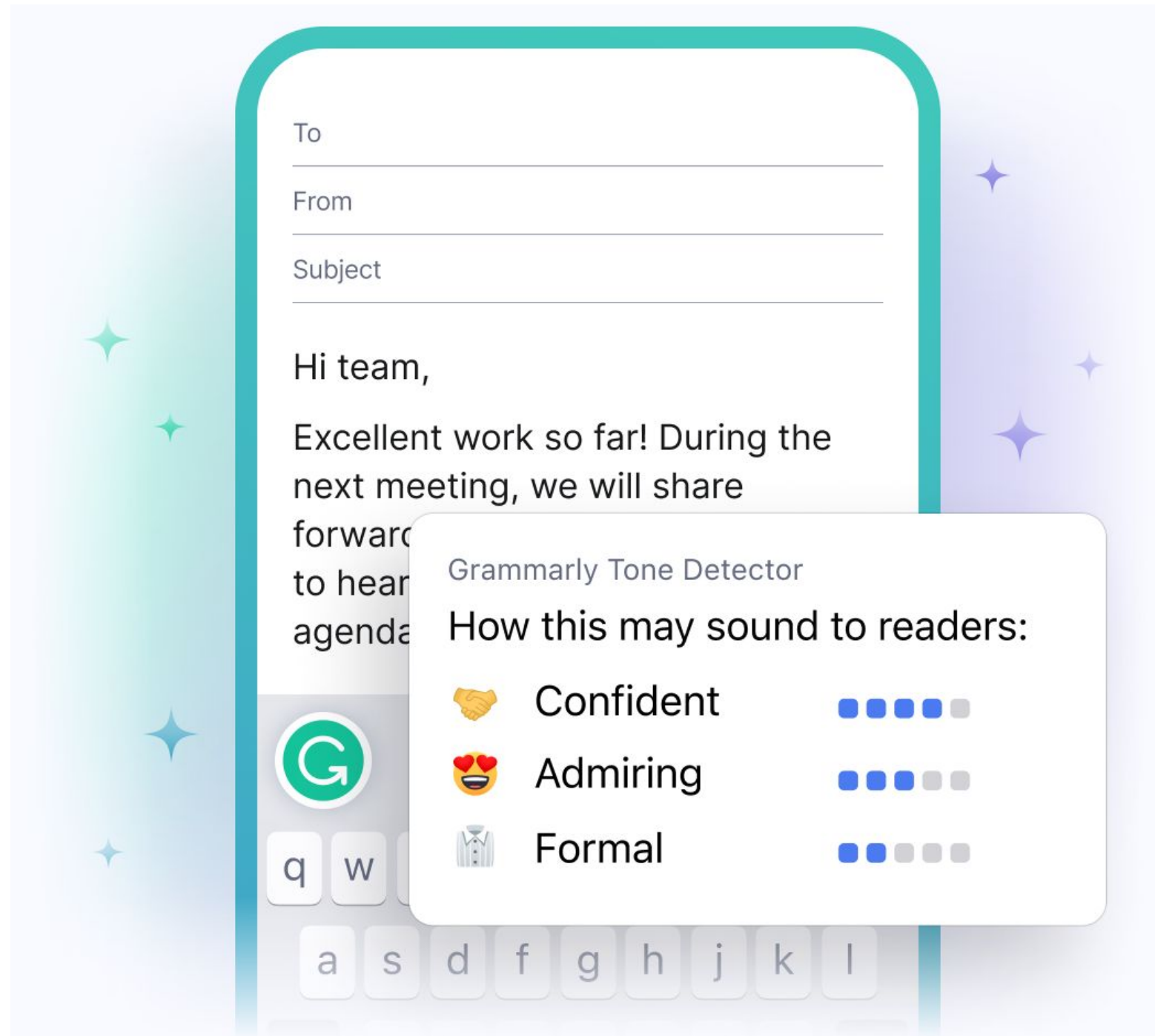
- ✓ Completeness
- ✓ Word Choice
- ✓ Typos

3+ more

Files we can read: DOC, DOCX, PDF, HTML, RTF, TXT

- Commonly used to process countless applicants' CVs
- CVs are parsed/scanned
- ATS weeds out CVs without keywords pertaining to desired skills
- Oftentimes, AI powered

Think Grammarly



- Specific to InforOS & Developer Portal
- Ultimate goal: provide feedback similar to Grammarly

Current Steps

Research AI models

- Explore deep learning and NLP
- Understand how and why they work

Start building

- Use Jupyter Notebook
- Read files & store text

Goals:

- Enforce uniformity for existing and future content
- Update/Remove outdated content

Potential Logistics

- AI Model: combination of GenAI and Predicative AI
 - GenAI: generates the comments/feedback
 - Predicative AI: generate new rules/guidelines for potential scenarios
- Input: any .txt, .pdf, and .docx (possibly) files
 - File is converted and text is extracted
- Output: writing to a file, comments or a short list of edits to be made
 - Ambition: having the ability to markup the file, like a human would when reviewing documents

Potential Obstacles

- Learning curve (AI models, training, etc.)
- Time
- Difficult to define the rules and guidelines
- May not be enough test cases to train the model
 - Encountering new situations, how does the model adapt?
- Costs:
 - Maintenance
- Is this even worth it?

This week:

- Continue research and explore different models
- Focusing on Sentiment Analysis and Transformer Models (BERT)



Thank you!

Steps

Create a custom algorithm

- Content guidelines are used as “rules”

Train an AI model using this algorithm

- Use AI, Jupyter Notebook, or VSCode

Goals:

- Enforce uniformity for existing and future content
- Update/Remove outdated content