

Capstone Project Proposal



Ugur Uresin

Business Goals

Project Overview and Goal What is the industry problem you are trying to solve? Why use AI/ML in solving this task? Be as specific as you can when describing how AI/ML can provide value. For example, if you're labeling images, how will this help the business?	<p>The key problem is age prediction and face recognition using AI/ML technologies.</p> <p>In the security industry, the challenge is to determine the age of individuals accurately and efficiently. This is crucial for access control, age verification, and ensuring compliance with age-related regulations.</p> <p>AI/ML is particularly suited for this task due to their ability to process and learn from large sets of images, recognizing patterns and features that correspond to different age groups.</p>
Business Case Why is this an important problem to solve? Make a case for building this product in terms of its impact on recurring revenue, market share, customer happiness and/or other drivers of business success.	<p>Importance for Security Business: For security businesses, implementing an AI-driven age detection system can lead to several benefits:</p> <p>Enhanced Compliance and Legal Safeguarding: Ensures adherence to age-specific regulations, reducing legal risks.</p> <p>Improved Efficiency: Automates age verification processes, speeding up entry procedures and reducing manpower costs.</p> <p>Increased Accuracy: Reduces human error in age estimation, enhancing the reliability of security checks.</p> <p>Customer Trust and Satisfaction: Accurate and quick verification processes can enhance customer experience, thereby potentially increasing client retention and referrals.</p>

Application of AI/ML

What precise task will you use AI/ML to accomplish? What business outcome or objective will you achieve?

The primary task is using AI/ML for precise age classification of individuals based on facial recognition. This will bolster security measures, particularly in venues or events where age restrictions are enforced. The outcome is a more secure, efficient, and compliant operational environment.

Success Metrics

Success Metrics

What business metrics will you apply to determine the success of your product? Good metrics are clearly defined and easily measurable. Specify how you will establish a baseline value to provide a point of comparison.

Key metrics might include:

Accuracy Rate: The percentage of correctly estimated ages compared to manual checks.

Processing Time: Reduction in time taken for age verification.

Data

Data Acquisition

Where will you source your data from? What is the cost to acquire these data? Are there any personally identifying information (PII) or data sensitivity issues you will need to overcome? Will data become available on an ongoing basis, or will you acquire a large batch of data that will need to be refreshed?

The dataset was acquired from Kaggle, suitable for training an AI model in age prediction.

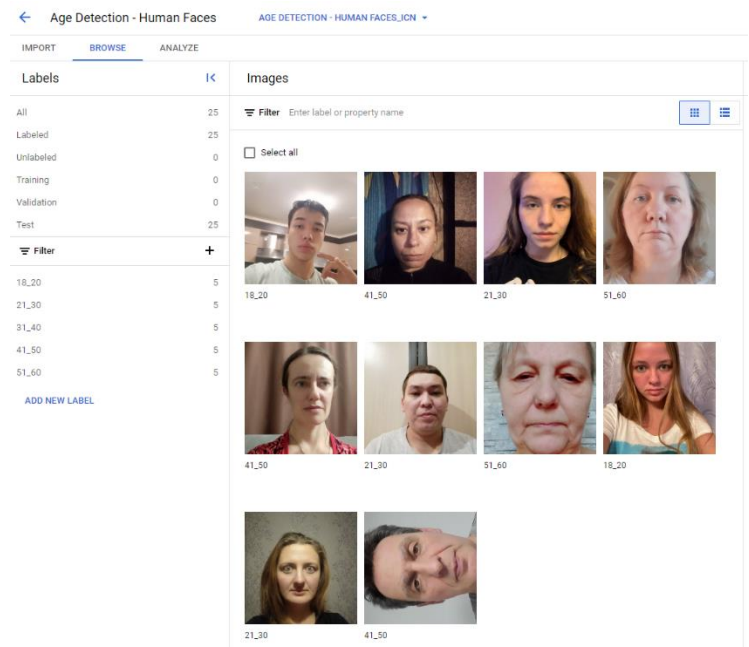
<https://www.kaggle.com/datasets/trainingdatapro/age-detection-human-faces-18-60-years>

The cost will include dataset acquisition, processing, and potentially ongoing costs for updates or expansions of the dataset. Considerations around PII are critical, especially in the security sector, necessitating strict adherence to data protection laws.

Data Source

Consider the size and source of your data; what biases are built into the data and how might the data be improved?

The dataset's size (150 images) might be a limiting factor, possibly not covering all demographic variations comprehensively. For a security-focused application, it's crucial to have a well-represented and diverse dataset to avoid biases that could lead to incorrect age predictions across different ethnicities and genders. Improvements could include expanding the dataset size and diversity to enhance the model's accuracy and reliability in real-world applications.



The full version of the dataset includes 95 000+ photos of people however it's not free and it's not purchased for this project.

Choice of Data Labels

What labels did you decide to add to your data? And why did you decide on these labels versus any other option?

Label definitions are given in the dataset. However, the labels are assigned to the images one by one in the Google Cloud's Vertex AI.

Labels	Images
18_20	30
21_30	30
31_40	30
41_50	30
51_60	30

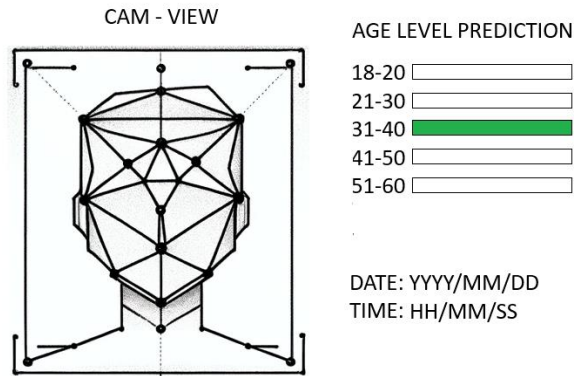
Model

Model Building How will you resource building the model that you need? Will you outsource model training and/or hosting to an external platform, or will you build the model using an in-house team, and why?	<p>The model is developed in Google Cloud's Vertex AI platform.</p> <p>Advantages:</p> <p>Control and Customization: In-house development allows to tailor the model precisely to your specific needs, especially important in the sensitive area of security.</p> <p>Data Security: Keeping the development process internal ensures better control over sensitive data, crucial in the age detection task for security purposes.</p> <p>Skill Development: The modeling team can develop and refine their skills in AI/ML, which is valuable for future projects.</p> <p>It also provides a chance to deliver the product much quicker compared to the from scratch development process.</p>
Evaluating Results Which model performance metrics are appropriate to measure the success of your model? What level of performance is required?	<p>Accuracy: This measures the percentage of correct predictions. In security contexts, accuracy above 95% is typically required.</p> <p>Precision and Recall: These metrics assess the ratio of correct predictions to total and actual positives. Both are expected to be high, often above 90%.</p> <p>F1-Score: A balance between precision and recall, with a high score (above 90%) indicating a well-balanced model.</p> <p>Confusion Matrix: Shows the model's performance for each class, with higher correct predictions (diagonal numbers) being ideal.</p> <p>In supporting officials with fast and reliable age verification, a satisfactory level of performance in key metrics is deemed sufficient, balancing accuracy with practicality.</p>

Minimum Viable Product (MVP)

Design

What does your minimum viable product look like? Include sketches of your product.



The user has opted for predicting age levels as a "single label". Given that this system is intended to serve as a decision support tool rather than the final authority, a simpler approach is favored.

To ensure no breach of personal data, no real photographs have been included in the sketch.

Use Cases

What persona are you designing for? Can you describe the major epic-level use cases your product addresses? How will users access this product?

Major Epic-Level Use Cases:

Security Checkpoints: Quick age verification at security checkpoints, essential in airports, government buildings, or age-restricted venues.

Regulatory Compliance: Ensuring adherence to age-specific legal requirements, particularly in controlled environments.

Accessing the Product:

Integrated Systems: The product will be accessible through systems integrated into existing security infrastructure, like turnstiles or entrance gates.

User Interface: A straightforward, easy-to-use interface designed for quick operation with minimal training.

Roll-out

How will this be adopted? What does the go-to-market plan look like?

To successfully introduce the AI-based age verification system to the security industry, the go-to-market plan includes:

Market Research:

Identify target segments in the security industry.

Understand specific needs for age verification.

	<p>Product Development: Collaborate with security experts for relevant features. Conduct pilot testing for real-world feedback.</p> <p>Partnerships: Form alliances with security firms and technology partners.</p> <p>Marketing and Sales: Develop educational and promotional materials. Participate in industry events and direct sales efforts. Establish channel partnerships for wider distribution.</p> <p>Customer Support: Offer comprehensive training and reliable support.</p> <p>Continuous Improvement: Regularly update the product based on user feedback.</p> <p>This streamlined approach focuses on targeted development, strategic partnerships, effective marketing, and ongoing support and improvement to ensure successful adoption in the security sector.</p>
--	---

Post-MVP-Deployment

<p>Designing for Longevity</p> <p>How might you improve your product in the long-term? How might real-world data be different from the training data? How will your product learn from new data? How might you employ A/B testing to improve your product?</p>	<p>Long-Term Product Improvement:</p> <ul style="list-style-type: none"> • Regularly update AI algorithms. • Continuously expand and diversify the dataset. • Integrate user feedback for enhancements. <p>Real-World Data vs. Training Data:</p> <ul style="list-style-type: none"> • May encounter unexpected variations and scenarios not covered in training data. <p>Learning from New Data:</p> <ul style="list-style-type: none"> • Implement periodic retraining with updated datasets. • Use techniques like online learning to adapt to new data trends. <p>Employing A/B Testing:</p> <ul style="list-style-type: none"> • Test new features or algorithms with a subset of users. • Evaluate different user interfaces and operational policies.
---	--

Monitor Bias

How do you plan to monitor or mitigate unwanted bias in your model?

Mitigating “Unwanted Bias”:

- Regularly review and update the training dataset for diversity.
- Use bias detection algorithms.
- Monitor model predictions for signs of bias and adjust accordingly.