****

**DataSpike (**[**https://www.dataspike.io/**](https://www.dataspike.io/)**)** is an independent international RegTech startup founded in 2020 and currently located in Dubai, UAE. We create AI/ML-based solutions to help our clients to detect entities or organizations partnership with whom may be risky for their business. Moreover, DataSpike has a unique Pay-As-You-Go pricing model, so clients do not need to commit to a year contract to screen their customers simply.

Features: Sanctions, PEPs and SIPs screening, Personal Watch lists, Adverse media, Fuzzy search by name and date of birth.

We are actively growing globally and seeking an experienced **Senior Machine Learning Engineer** to lead computer vision and deep learning technology development.

**What would you do:**

* Developing and implementing algorithms for image forgery detection: This may involve researching and experimenting with different methods for identifying and detecting manipulated images, such as image analysis, machine learning, and computer vision techniques
* Collecting and preprocessing data. This includes gathering large datasets, performing data labeling, and preparing the data for analysis by cleaning, resizing, and formatting it appropriately
* Training and testing the models using statistical methods and machine learning algorithms to train models on the collected data, and testing the performance of the models using various evaluation metrics
* Analyzing and interpreting results
* Collaborating with other teams such as IT, software development, and legal, to ensure that the image forgery detection methods are implemented in a manner that respects data privacy and security
* Keeping updated with the recent research in the field of image forgery detection

**Key Qualifications:**

* BS, MS, or PhD Degree in Computer Science/Machine Learning or similar
* 5+ years of experience in ML/ CV area
* Fraud detection/image forgery domain expertise as an advantage
* Strong background in Python
* Proactivity, responsibility for own initiatives
* Ability to collaborate effectively with teammates
* Being able to complete end to end delivery cycle from research and up to solution running on production

**Assignment:**

Document Forgery Detection Mini-Project

**Task:**

Create a simple model for Image Forgery Detection. The solution must be as fast as possible, ideally up to 1 second per image.

If you choose to train a neural network, please use either PyTorch or TensorFlow framework.

**Instructions:**

Data Preparation:

Download the [MICC\_F600] (<http://lci.micc.unifi.it/labd/2015/01/copy-move-forgery-detection-and-localization/>) dataset (or any other similar dataset you prefer).

Apply image augmentations, if necessary.

Model Building:

Build a simple model/algorithm.

The model should take in images and output whether the input image is genuine or forged.

You can experiment with network architectures, layers, and hyperparameters.

Indicate, why you chose this particular algorithm and what methods you considered.

Training and Evaluation:

Split the dataset into training and testing sets.

Train the model on the training set.

Evaluate the model's performance on the testing set.

Calculate relevant metrics of model performance.

**Deliverables:**

A Jupyter/Colab Notebook or Python script containing your code.

Documented explanations for each step in the code.

Plots or visualizations showcasing model performance metrics.

(Optional) Additional insights or improvements you've implemented.

**Evaluation Criteria:**

Demonstrated critical thinking in problem-solving.

Proper dataset preprocessing.

Effective model building and training.

Accurate evaluation of model performance metrics.