

Job Description for Digilytics™ Data Scientist - Document Extraction

| | |
|----------------------|--------------------------------------|
| Role | Data Scientist – Document Extraction |
| Position Type | Full Time |
| Job Type | Data Science (min 3 yrs) |

About Us

At Digilytics™, we build and deliver easy to use AI products to the secured lending and consumer industry sectors. In an ever-crowded world of clever technology solutions looking for a problem to solve, our solutions start with a keen understanding of what creates and what destroys value in our clients' business.

Founded in 2014, by Arindom Basu, the leadership of Digilytics™ is deeply rooted in leveraging disruptive technology to drive profitable business growth. With over 50 years of combined experience in technology-enabled change, the Digilytics™ leadership is focused on building a values-first firm that will stand the test of time.

We are currently focussed on developing a product, Revel FS, to revolutionise loan origination for mortgages and secured lending. We are also developing a second product, Revel CI, focused on improving trade (secondary) sales to consumer industry clients like auto and FMCG players.

The leadership strongly believes in the ethos of enabling intelligence across the organization. Digilytics AI is headquartered in London, with presence across India.

About the role

Digilytics is pioneering the development of intelligent mortgage solutions in International and Indian markets.

We are looking for experienced data scientists, who have the aspirations and appetite for working in a start-up environment, and with relevant industry experience to make a significant contribution to our Digilytics™ platform and solutions. Primary focus would be to apply machine learning techniques for data extraction from documents from variety of formats including scans and handwritten documents.

Responsibilities

- Develop a learning model for high accuracy extraction and validation of documents, e.g. in mortgage industry
- Work with state-of-the-art language modelling approaches such as BERT combining all three AI streams – NLP, computer vision and machine learning
- Understand the Digilytics™ vision and help in creating and maintaining a development roadmap
- Interact with clients and other team members to understand client-specific requirements of the platform
- Contribute to platform development team and deliver platform releases in a timely manner

- Liaise with multiple stakeholders and coordinate with our onshore and offshore entities
- Evaluate and compile the required training datasets from internal and public sources and contribute to the data pre-processing phase.

Expected and Desired Skills

- Either of the following Deep learning frameworks – PyTorch (preferred) or Tensorflow
- Ability to evaluate competing Neural Network architectures such as Graph Neural Networks, Transformer Networks, etc
- Python used for analytics applications including data pre-processing, EDA, statistical analysis, machine learning model performance evaluation and benchmarking
- Good scripting and programming skills to integrate with other external applications
- Good interpersonal skills and the ability to communicate and explain models
- Ability to work in unfamiliar business areas and to use your skills to create solutions
- Ability to both work in and lead a team and to deliver and accept peer review
- Flexible approach to working environment and hours

Experience

- Between 3-5 years of relevant experience
- Hands-on experience with
 - Python and/or R
 - Machine Learning
 - Deep Learning (desirable)
- End to End development of a Deep Learning based model covering model selection, data preparation, training, hyper-parameter optimization, evaluation, and performance reporting.
- Proven experience working in both smaller and larger organisations having multicultural exposure
- Domain and industry experience by serving customers in one or more of these industries - Financial Services, Professional Services, other Consumer Industries

Education Background

- A Bachelor's degree in the fields of study such as Computer Science, Mathematics, Statistics, and Data Science with strong programming content from a leading institute
- An advanced degree such as a Master's or PhD is an advantage