# Data Scientist
# Pratham Digital

## About Pratham

Pratham was founded in 1995, to provide pre-school education to children in Mumbai slums. Over the last 20 years Pratham has grown to be India's largest NGO working to provide quality education to underprivileged youth and children in over 21 states and union territories across the country, with a range of interventions.

Pratham is a widely recognized organization, having received notable awards such as the WISE Prize for Innovation, Skoll Award for Social Entrepreneurship, the Henry R Kravis Prize in Leadership and the CNN-IBN Indian of the Year for Public Service. For more details, refer to www.pratham.org

## About Pratham Digital

Pratham started its digital intervention with the Hybrid learning program in 400 villages of Rajasthan, Maharashtra and Uttar Pradesh in the year 2015. In 2017 with the support of Google.org, and Sarva Mangal Family Trust this program expanded to over 1000 villages. The support led to the formation of core groups within Pratham which produced over 350 videos in and about 70 learning games and software needed to deploy and monitor digital resources in the village communities. These resources are present in 10 regional languages and English

Subsequently, the digital resources were also made available in Pratham's foundational learning camp programs and also in the Early Childhood Education support program on an experimental basis. The digital learning material (games and videos) created for different age groups is available on Google Playstore as the PraDigi app, which was launched in October 2017 along with Youtube and other learning platforms.

The digital hardware and software are currently available in various Pratham programs **across 21 states with content in 11 languages** including Punjabi, Assamese, Bengali, Odiya, Telugu, Tamil, Kannada, Marathi, Gujarati, Hindi and English. The games are developed in HTML5/JavaScript on that they can be embedded on web pages for an online version or used on desktops in an offline version.

## Data Scientist – Job Description

We're looking for talented people who will put our goal to develop innovative educational methodologies at the center of everything we do.

We need a data scientist who will help us discover the information hidden in vast amounts of data that we have collected over the years, and help us make smarter decisions to deliver even better products and content. Your primary focus will be in applying data mining techniques, doing statistical analysis, and building high quality automated assessment tool using machine learning techniques.

## Responsibilities

- Think creatively and identify opportunities to leverage machine learning in order to improve a learner's learning experience
- Creating automated student assessment system and constant tracking of its performance
- Use advanced technologies such as speech and vision synthesis to evaluate non textual response to questions for assessing soft skills.

- Ability to use NLTK and identify words related to the defined keywords would be critical
- Ability to use NLP to provide feedback on learner response
- Translate and summarize complex analysis into understandable, actionable insights and recommendations that directly drive effective content delivery strategy
- Data mining using state-of-the-art methods
- Develop machine learning and other AI models with Python, R, or other languages and tools
- Enhancing data collection procedures to include information that is relevant for building analytic systems
- Processing, cleansing, and verifying the integrity of data used for analysis
- Work effectively in a team environment, as well as independently, to deliver against key initiatives
- Take initiatives and drive each project to completion with minimal guidance while effectively managing multiple projects at a time
- Contribute to a positive and supportive team culture.
- Work closely with our software engineers to put algorithms into practice
- Mentor and provide direction to other members in the team.

## Desired Qualifications and Experience

### Required:
- Bachelors in mathematics, statistics, engineering or computer science or related field; Masters or PHD degree preferred.
- 5+ years of relevant quantitative and qualitative research and analytics experience.
- Extensive knowledge and practical experience in several of the following areas: machine learning, statistics, NLP, deep learning, recommendation systems, dialogue systems, information retrieval
- Skilled with Java, C++, or other programming language, as well as with R, MATLAB, Python or similar scripting language
- Experience with common NLP techniques, such as Pre-processing (tokenization, part-of-speech tagging, parsing, stemming); Semantic analysis (named entity recognition, sentiment analysis); Modeling and word representations (TF-IDF, LSA, LDA, word2vec)
- Excellent understanding of machine learning techniques and algorithms, such as k-NN, Naive Bayes, SVM, Decision Forests, etc.
- Experience with data visualisation tools, such as Power BI
- Proficiency in using query languages such as SQL
- Experience with NoSQL databases, such as Datastore
- Ability to articulate the strengths and weaknesses of various predictive modeling techniques
- Strong understanding of statistical testing necessary to assess model performance
- Great communication skills and ability to generate discussions around data analytics
- Inquisitive mind and willingness to make the difference
- Excellent track record of original research is highly desirable

## Application Process

Send the following to digital@pratham.org (early applicants will be given preference) and mention 'Application for the position of Data Scientist'in the subject line, with the following attachments:

1. **Current Résumé**: Résumé should contain:
   - Contact Information for Applicant
   - Academic Background, universities attended/degrees acquired
   - Past work experience, highlighting relevant skills
   - Languages Spoken