

## Data Science Internship JD

---

This role will be engaged in data science-related research and software application development and engineering duties related to our enterprise-grade Wi-Fi technology and autonomous platform to provide an unprecedented visibility into the user experience. The Software Engineer will collaborate with other engineers and product managers to build the next generation of autonomous Wi-Fi networks leveraging big data and predictive models. This position requires experience in dealing with huge amounts of data generated by communication protocols. The Software Engineer will use his/her knowledge of wireless communication networks, machine learning and software engineering to develop and implement scalable algorithms to process a large amount of streaming data to detect anomalies, predict problems, and classify them in real-time. In addition, the Software Engineer will leverage the data collected from the Wi-Fi network to empower the inference engine of our Mist platform and systems, including the Mist virtual assistant chat bot. The Software Engineer will also be responsible for determining the likelihood of failures across the Wi-Fi network and performing failure scope analysis.

### Responsibilities

- Utilize analytical and programming skills and open-source systems, such as Apache Storm, Apache Spark, Elasticsearch, Redis, etc. develop data processing pipeline required efficacy and latency
- Design and implement machine learning solutions which require to process terabytes of streaming data to detect anomalies in Wi-Fi networks of our customers, predict problems and future trends, classify them in real-time.
- Analyse and organize raw data in JSON or Protobuf format, build data systems and pipelines and evaluate business needs and objectives
- Troubleshoot production environment and customer reported issues.

# Data Science Internship JD

---

## Qualifications

- Proficient in Java and python programming languages
  - Good knowledge and experience of the big data tool sets and techniques of distributed storage and computation engine
  - Experience to develop the reusable and highly scalable data processing component.
  - Good knowledge and experience to work with cloud based CICD tools and cloud devops teams to collect stats and create monitors for our data processing pipelines
  - Knowledge and experience of the multi-cloud production environment
  - Require the agility to troubleshoot open-source data processing engine, such as Apache Spark and Apache Storm
-