

Hello, Truecaller is calling you from Bangalore, India! Ready to pick up?

Our goal is to make communication smarter, safer, and more efficient, all while building trust everywhere. We're all about bringing you smart services with a big social impact, keeping you safe from fraud, harassment, scam calls or messages, so you can focus on the conversations that matter.

- Top 20 most downloaded apps globally, and world's #1 caller ID and spam-blocking service for Android and iOS, with extensive AI capabilities, with more than **400 million** active users per month.
- Founded in 2009, listed on Nasdaq OMX Stockholm and is categorized as a Large Cap. Our focus on innovation, operational excellence, sustainable growth, and collaboration has resulted in consistently high profitability and strong EBITDA margins.
- A team of 400 people from ~35 different nationalities spread across our headquarters in Stockholm and offices in Bangalore, Mumbai, Gurgaon and Tel Aviv **with high ambitions**.

We in the Insights Team are responsible for SMS Categorization, Fraud detection and other Smart SMS features within the Truecaller app. The OTP & bank notifications, bill & travel reminder alerts are some examples of the Smart SMS features. The team has developed a patented offline text parser that powers all these features and the team is also exploring cutting edge technologies like LLM to enhance the Smart SMS features. The team's mission is to become the World's most loved and trusted SMS app which is aligned with Truecaller's vision to make communication safe and efficient. Smart SMS is used by over 90M users every day.

As a Senior Data Scientist, you will be responsible for collecting, organizing, analyzing, and interpreting Truecaller data with a focus on NLP. In this role, you will be pivotal in advancing our work with large language models and on-device models across diverse regions. Your expertise will enhance our natural language processing, machine learning, and predictive analytics capabilities.

What you bring in:

- 5+ years of experience in designing, developing, and deploying ML models at scale, with a focus on NLP-driven solutions.
- Strong background in Natural Language Processing (NLP), including text classification, entity recognition, language modeling, and transformer-based architectures.

- Experience in building and deploying models at scale, handling millions of messages efficiently while maintaining performance and accuracy. Also working with on-device models.
- Ability to not only build ML models but also take ownership of deploying them into production, ensuring scalability, reliability, and monitoring.
- Knowledge of anomaly detection, adversarial ML techniques, and risk modeling to identify and prevent spam and fraudulent messaging activities.
- Strong ability to take ML models from research and experimentation to production, working closely with ML engineers and data engineers.
- Expertise in machine learning libraries such as TensorFlow, PyTorch, pandas and Scikit-learn, along with NLP-specific tools like Hugging Face Transformers, spaCy with experience in TFLite, ONNX.
- Hands-on experience fine-tuning LLMs including transformer-based architectures (BERT, GPT, LLaMA, T5, etc.) for domain-specific applications, including knowledge distillation, quantization, and model compression for efficiency.
- Strong ability to design, refine, and optimize prompts for LLM-based applications, ensuring high-quality responses and reduced model hallucinations.
- Ability to leverage data driven decision by experimentation, and statistical analysis to improve models and business outcomes.
- Strong understanding of designing, testing, and optimizing prompts for LLM-based applications to improve model accuracy and efficiency.
- Programming knowledge in at least one language, such as Python or R. Preferably python.
- Expert knowledge of machine learning algorithms.
- Familiarity with database modelling and data warehousing principles with a working knowledge of SQL
- Experience in building and optimizing large-scale data processing systems using Spark/PySpark
- Strong ability to work cross-functionally with engineers, product managers, and business stakeholders to align ML solutions with company objectives.

The impact you will create:

- Take a loosely defined business problem and break it into tractable data problems. For each data problem, clearly articulate the value of solving it, its impact, and its complexity.

- Collaborate with Product and Engineering to scope, design, and implement systems that solve complex business problems ensuring they are delivered on time and within scope.
- Design, develop, and optimize state-of-the-art NLP models for large-scale message classification, fraud detection, and spam filtering, impacting millions of users globally.
- Take full ownership of ML model development, deployment, and monitoring, ensuring models are production-ready, scalable, and cost-efficient.
- Lead data science projects from ideation to deployment, ensuring alignment with business objectives and timelines.
- Manage and analyze large datasets collected from multiple countries, ensuring data integrity and consistency.
- Stay updated on industry best practices and emerging technologies to drive innovation within the Data Team.
- You work collaboratively across systems and teams to solve user and business problems. You are expected to help define success and design and build the systems to achieve it.
- To work with the Product to decide on priorities and set direction, design solutions, and help the team implement them.

It would be great if you also have:

- Understanding of Conversational AI
- Deploying NLP models in production
- Working knowledge of GCP components

Life at Truecaller – Behind the code: <https://www.instagram.com/lifeattruecaller/>

Sounds like your dream job?

We will fill the position as soon as we find the right candidate, so please send your application as soon as possible. As part of the recruitment process, we will conduct a background check.

This position is based in Bangalore, India.

We only accept applications **in English**.

What we offer:

- **A smart, talented and agile team:** An international team where ~35 nationalities are working together in several locations and time zones with a learning, sharing and fun environment.

- **A great compensation package:** Competitive salary, 30 days of paid vacation, flexible working hours, private health insurance, parental leave, telephone bill reimbursement, Udemy membership to keep learning and improving and Wellness allowance.
- **Great tech tools:** Pick the computer and phone that you fancy the most within our budget ranges.
- **Office life:** We strongly believe in the in-person collaboration and follow an office-first approach while offering some flexibility. Enjoy your days with great colleagues with loads of good stuff to learn from, daily lunch and breakfast and a wide range of healthy snacks and beverages. In addition, every now and then check out the playroom for a fun break or join our exciting parties and or team activities such as Lab days, sports meetups etc. There something for everyone!

Come as you are: Truecaller is diverse, equal and inclusive. We need a wide variety of backgrounds, perspectives, beliefs and experiences in order to keep building our great products. No matter where you are based, which language you speak, your accent, race, religion, color, nationality, gender, sexual orientation, age, marital status, etc. **All those things make you who you are, and that's why we would love to meet you.**