AI Engineer

## **About the job**

Join our mission

Nebuly AI is a well-funded, fast-growing VC-backed startup in the emerging AI sector. Nebuly is not only pioneering but defining a new market category: LLMs user analytics. Recognizing the immense potential of Generative AI, we are ambitiously charting unexplored territories, crafting a unique niche that no company has ventured into before.

Thanks to Nebuly, companies can automatically capture how users interact with their LLM-powered chatbots / agents. Our platform is key to understand what LLM users like, what they don’t and why, what are the most asked questions and how you can create personalised LLMs responses to delight your customers.

Despite our youthful journey, demand for our product is skyrocketing, and our robust funding ensures we're poised for strong growth. Our team is a collective of forward-thinkers, committed to transforming industries with revolutionary AI products.

We offer not only very competitive salaries, but also the opportunity to shape the AI landscape.

If you’re a skilled Python engineer with strong knowledge of AI systems, we want you.

Examples of your daily responsibilities

As an AI Engineer, you will oversee data labeling processes and ensure the quality and consistency of training datasets critical to model performance. You will fine-tune proprietary large language models (LLMs) utilizing cloud-based GPU infrastructure, optimizing model accuracy and efficiency for specific business applications. Additionally, you’ll be responsible for integrating model outputs seamlessly into the broader platform, collaborating closely with software engineers to ensure robust deployment and optimal functionality. Your responsibilities will include:

* Oversight data labeling and training data quality
* Fine-tune proprietary LLM models on cloud GPUs
* Integrate the model outputs with the rest of the platform
* Oversight on model output quality with both automatic tools and human supervision
* Develop PoCs for AI features and once completed integrate them into production

What we look for

* Minimum of 3 years of professional experience in training and evaluating Neural Networks, with a particular emphasis on NLP tasks.
* Experience in training large deep learning models on multiple GPUs
* Professional experience in using Python and common deep learning frameworks as PyTorch, HuggingFace Transformers and vLLM.
* Strong knowledge and experience with microservices architecture and building scalable systems.
* Strong knowledge of PostgresSQL and experience in working with ORMs like SQLAlchemy.
* Proficiency in designing and implementing RESTful APIs.
* Strong knowledge and experience with testing practices and frameworks, such as unit testing, integration testing, and end-to-end testing, using libraries like Pytest.
* Understanding of DevOps practices and CI/CD pipelines.

What we offer

High-impact work in small, fast-moving teams. At Nebuly, you’ll work in small, entrepreneurial teams with a high degree of ownership and autonomy. Regardless of your seniority, your contributions will have a direct impact—from the earliest ideas to product launch. You’ll have the chance to build things from scratch and see your code evolve into real products used at scale.

A platform for real growth. As more and more companies rely on our platform to understand and optimize their AI experiences, new challenges and opportunities emerge constantly. This means your role won’t stay static—you’ll keep growing with the product, facing new technical and strategic problems as we scale.

Competitive compensation & stock options. We offer a competitive salary, tailored to your experience and location. In addition, you’ll have access to stock options so you can share in the value we’re creating together.

Remote-friendly & flexible. We believe great work doesn’t have to happen in a fixed place or on a fixed schedule. We’re remote-friendly, offer flexible working hours, and support you in setting up the conditions to do your best work—wherever you are.

The selection process

For most roles at Nebuly, here’s what our hiring process typically looks like:

* Introductory Call (30min): You’ll first speak with the person who initially contacted you or the hiring coordinator for the role. This is an informal chat where we’d love to learn more about you—what brought you to apply, your interests, and what excites you. We were impressed by your profile and want to give you space to ask questions and get a better sense of what we do at Nebuly. 🗣 *All conversation are held in English.*
* Conversation with Your Future Manager (30min): Next, you’ll meet the person you’d be directly reporting to. This call dives deeper into the role: the day-to-day responsibilities, challenges, and what success looks like. For technical roles, *this will include a practical or coding interview*.
* Logic Challenge (30min): A test with logic-based questions. This helps us understand how you approach unfamiliar problems.
* Behavioral Test (30min): You’ll complete a short online behavioral assessment. This isn’t about right or wrong answers—it’s designed to give us insight into your working style, how you collaborate, and what kind of environment helps you do your best work. It also helps us ensure strong alignment between you and the culture at Nebuly.
* Meeting the Co-founders (30min): If you’ve reached this step, things are looking great. You’ll meet Nebuly’s co-founders to talk about your motivations, career goals, and what makes you thrive at work. It’s also your chance to get to know the people shaping the vision of the company.
* Role-Focused Project (2 hrs): At this stage, we’re genuinely excited about the possibility of working together. We’ll propose a short project that reflects the kind of work you’d actually be doing at Nebuly. You’ll have full flexibility on timing—we’ll work around your schedule to make sure it’s manageable and enjoyable.
* Deep-Dive Technical Session (1hr): This is a second round of technical questions, typically with another member of the team or the lead of your function. It’s a chance for us to go deeper into your technical thinking and problem-solving approach, and for you to show how you tackle realistic challenges you’d face at Nebuly.
* Offer 🎉: If it’s a match, we’ll give you a call to walk through the offer details—compensation, starting date, next steps—and answer any questions you might have.

Please submit your resume at careers@nebuly.ai, along with any relevant code samples or GitHub repositories, showcasing your backend development skills.