Chapter 8
Assessing Security and Privacy Concerns in LLM Applications: Legal and Social Concerns

Bhupinder Singh
https://orcid.org/0009-0006-4779-2553
Sharda University, India

Arunima Shastri
https://orcid.org/0000-0002-6260-1253
Christ University, Bangalore, India

Saurabh Chandra
https://orcid.org/0000-0003-4172-9968
Bennett University, Greater Noida, India

ABSTRACT

Large language models (LLMs) like Open AI's GPT-4 have transformed many industries, including natural language processing. But, still the applications of these need a properly secure mechanism as it is getting attached to everyday technology at on an extensive rate and large security & privacy risk are encapsulated in its use which must be controlled for safe and ethical use. LLM models are trained with an avalanche of data, often from web-scraped sources that could be handling personal and sensitive content. This may inadvertently have the LAZ model hold on to, and possibly leak, some private data of users who use the API. Having measures to address them through comprehensive strategies and conforming with regulations are important in the responsible usage of LLM applications.

DOI: 10.4018/979-8-3693-8387-2.ch008

INTRODUCTION

Ensuring that training data are indeed well-vetted and fully anonymized is essential to avoid these risks. The input data which the users provide to an LLM can be sensitive in nature (Dinesh Arokia Raj et al., 2024). These interactions are managed by infrastructure, and a compromise of that result in data breaches (Yue & Shyu, 2024). To prevent unauthorized access, user inputs should be handled and stored securely. The primary step to secured LLM applications is universal encryption and secures data transmission s (Mithas et al., 2022). The LLMs can also be vulnerable to adversarial attacks, where they are supplied with the inputs which have been crafted in such a way that it changes its behavior (Ivanov et al., 2019). They can vary from input manipulations causing the model to generate harmful or misleading content, all the way up to extracting sensitive information for advanced attackers (Javaid et al., 2022). Strong security measures, such as vigilant monitoring and updating of models are essential to identify/adapt against adversarial activities like this (Fraga-Lamas et al., 2021).

Apart from the obvious technical security issues LLMs bring along a whole raft of ethical considerations - and specifically around bias, discrimination (Asadollahi-Yazdi et al., 2020). These models can both learn from previous data and thus inherits any biases that were present within the training data, as well as propagate this potential societal bias (Meyendorf et al., 2023). This results in outputs that perpetuate stereotypes or unjustly target certain groups (Zhong et al., 2017). This is something which needs continuous work and we will need to audit the models again to fine tune them for fairness, bias reduction (Angelopoulos et al., 2019). That is why it must be regulated to protect the privacy and security of consumers, as more LLMs are used (Lu et al., 2020). For LLM applications, regulatory compliance in the areas of data protection such as General Data Protection Regulation (GDPR) is a must (Chander et al., 2022). These laws require specific data care and transparency standards so that users know when their information is used or stored (Kasowaki & Ahmet, 2024). Though LLMs and Transformers of this sort do wonderful tasks, they introduce really large security and privacy issues (Sima et al., 2020). Tempted with an example, a model trained on not anonymized datasets could generate responses that contain private information thus becoming the privacy time bomb. But as technology changes, the challenge evolves and vigilance in perpetuity appears to be a necessity if we are going to protect the privacy of being tracked by these powerful tools (Tseng et al., 2021).