Journal of Global Economy, Business and Finance (JGEBF) ISSN: 2141-5595
DOI: 10.53469/jgebf.2025.07(09).02

AI in Regulatory Compliance: Revolutionizing Training with Fine-Tuned Language Models for Banking and Payment Systems

Chandre Gowda

Senior Lead Software Engineer at a Fortune 500 Bank, Richmond, Virginia, United States

Abstract: ***This article examines the use of fine tuned small Language Learning Models LLMs to develop AIdriven regulatory compliance training tailored for banking and payment systems. It discusses the integration of industry specific regulations, internal policies, and best practices into training modules, with a focus on personalized learning, interactive simulations, and real time updates. The proposed solution aims to improve training efficiency, ensure adherence to regulations, and deliver scalable, cost effective compliance training solutions. Additionally, the article addresses potential challenges, including data quality, model interpretability, and ethical considerations in AI deployment for regulatory training.***

Keywords: Regulatory compliance, Language Learning Models (LLMs), AI - driven training, Banking, Payment systems, Fine - tuning, Personalized learning, Interactive modules, Real - time updates

1. Introduction

Regulatory compliance in banking and payment systems is complex, requiring constant updates and employee training. Traditional training methods are often static and cannot keep pace with regulatory changes. This white paper proposes leveraging fine - tuned small LLMs to generate customized, up - to - date training materials tailored to specific roles within the bank, thereby ensuring comprehensive and effective compliance training.

2. Technical Architecture and Training Process

2.1 Model Architecture

The proposed solution uses a transformer - based architecture, similar to GPT (Generative Pre - trained Transformer) models but scaled down to a more manageable size for fine - tuning and deployment within organizational infrastructure [7]. The model consists of:

- Embedding layer: Converts input tokens into continuous vector representations
- Multiple transformer blocks: Each containing self - attention and feed - forward neural network layers
- Layer normalization and residual connections: To stabilize training and improve gradient flow
- Output layer: Generates probability distributions over the vocabulary for next token prediction

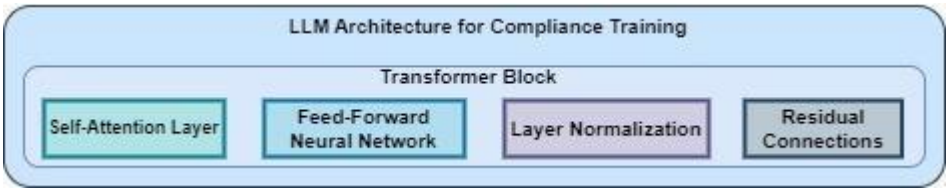

Figure 1: Illustrates the overall architecture of the fine - tuned LLM for compliance training

2.2 Fine - Tuning Process

The fine - tuning process adapts the pre - trained LLM to the specific domain of regulatory compliance [8]. Figure 2 depicts the steps involved:

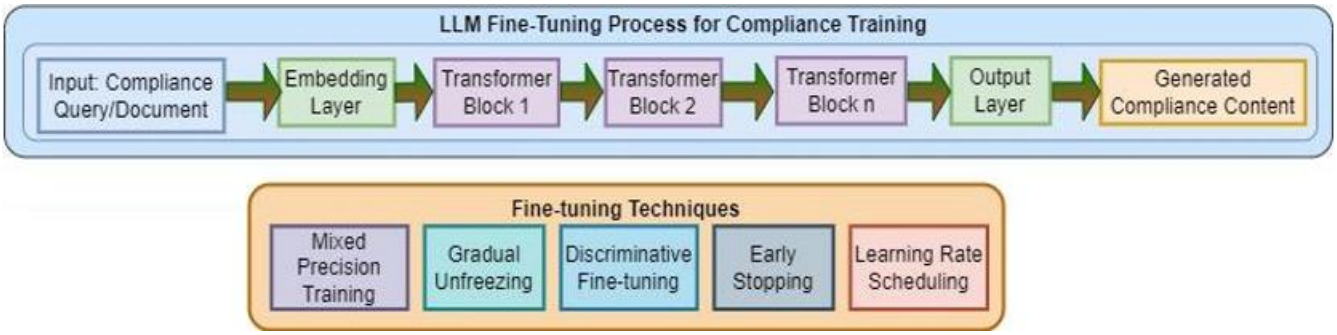

Figure 2: Illustrates the fine tuning process for Compliance Training and Fine - Tuning Techniques

1) Pre - training: Start with a small LLM pre - trained on a large corpus of general text data.
2) Domain - specific data collection: Gather a comprehensive dataset of regulatory documents, internal policies, and compliance guidelines.
3) Data preprocessing: Clean and format the collected data, ensuring consistency and relevance.
4) Fine - tuning:
   - Use techniques like gradual unfreezing and discriminative fine - tuning to adapt the model efficiently.
   - Mixed precision training is employed to reduce memory usage and enhance training speed.
   - Implement early stopping and learning rate scheduling to prevent overfitting.
5) Evaluation: Assess the model's performance using domain - specific metrics and expert review.
6) Iterative refinement: Continuously update the model with new regulatory information and feedback.

2.3 **Deployment and Integration**

- Deploy the fine - tuned model on - premises or in a secure cloud environment.
- Implement an API layer for seamless integration with existing learning management systems.
- Establish a pipeline for regular model updates and version control.

3. Key Features and AI Strategies

3.1 Personalized Learning Paths

- Utilize collaborative filtering and content - based recommendation systems to suggest relevant training modules.
- Implement adaptive testing algorithms to adjust difficulty based on user performance.

3.2 Interactive Learning Modules

- Develop reinforcement learning - based simulations for complex compliance scenarios.
- Use natural language generation to create dynamic quiz questions and explanations.

3.3 Real - Time Updates

- Implement named entity recognition and topic modeling to automatically categorize new regulatory documents.
- Use text summarization techniques to generate concise updates for employees.

3.4 AI - Powered Assistance

- Deploy a transformer - based question - answering system to deliver on - demand compliance guidance [9].
- Implement intent classification and entity extraction for accurate query interpretation.

3.5 Hybrid Approach

AI - Assisted Content Generation with Human Oversight

While the fine - tuned LLM can generate high - quality compliance training content, a hybrid approach combining AI -
generated content with human expert review ensures the highest standards of accuracy and relevance. This approach is particularly useful for generating training questions and assessments. Figure 3 illustrates this process:

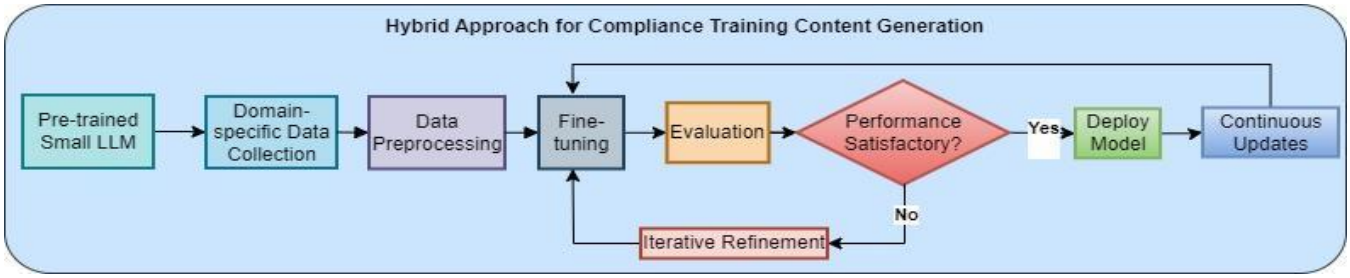

Figure 2: Illustrates the Hybrid Approach of AI - Assisted Content Generation with Human Oversight

The hybrid approach works as follows:

AI - Generated Draft Questions: The fine - tuned LLM generates a set of draft questions based on the compliance topics and regulations it has been trained on. These questions

cover various aspects of regulatory compliance, including multiple - choice, true/false, and open - ended formats.

Human Expert Review: Compliance experts review the AI - generated questions, assessing them for accuracy, relevance, and appropriateness. The experts can:

- Approve questions that meet the required standards
- Revise questions that are close but need refinement
- Reject questions that are unsuitable or irrelevant

Iterative Refinement: Revised questions can be fed back into the LLM for further improvement, creating a feedback loop that enhances the model's performance over time.
Final Question Bank: Approved questions are added to the final question bank, which forms the basis of compliance training assessments and quizzes.

Benefits of this hybrid approach include:

- **Efficiency**: The AI significantly reduces the initial effort required to generate a large pool of diverse questions.
- **Quality Assurance**: Human oversight ensures that all questions adhered to regulatory standards and organizational requirements.
- **Continuous Improvement**: The feedback loop helps the LLM learn from expert revisions, improving its question generation capabilities over time.
- **Customization**: Experts can tailor questions to specific roles, departments, or regulatory focus areas.
- **Scalability**: This approach allows for the rapid creation and updating of training materials in response to new regulations or organizational changes.

Implementation Considerations:

- **Expert Interface**: Create a user friendly interface for compliance experts to review, edit, and manage AI generated questions efficiently.
- **Feedback Mechanism**: Implement a system to capture expert feedback and use it to fine - tune the LLM further.
- **Version Control**: Maintain a versioning system for questions to track changes and updates over time.
- **Audit Trail**: Keep a record of AI - generated content and subsequent human modifications for transparency and regulatory compliance.
- **Performance Metrics**: Establish metrics to measure the effectiveness of the hybrid approach, such as the percentage of AI - generated questions approved without revision and the time saved compared to traditional content creation methods.

By combining the strengths of AI generated content with human expertise, this hybrid approach provides an effective solution for creating high quality, uptodate compliance training materials efficiently.

4. Case Studies

4.1 JPMorgan Chase: AI - Driven AML Compliance

JPMorgan Chase implemented an AI - driven anti - money laundering (AML) system that reduced false positives by 80% and increased the detection of suspicious activities by 20% [1]. While not specifically a training system, this case demonstrates the potential of AI in enhancing compliance processes.

4.2 HSBC: Machine Learning for Compliance Monitoring

HSBC deployed a machine learning system to monitor compliance - related communications, reducing the need for manual reviews by 70% and improving accuracy in identifying potential compliance issues [2]. This system indirectly supports ongoing compliance training by highlighting areas that require additional focus.

5. Limitations and Challenges

5.1 Data Quality and Availability

- Ensuring comprehensive and up - to - date training data can be challenging, especially for niche regulatory areas.
- Data privacy concerns may limit the availability of real - world examples for training.

5.2 Model Interpretability

- The "black box" nature of deep learning models can make it difficult to explain specific training recommendations or decisions.
- Regulatory bodies may require a certain level of interpretability for compliance - related systems. [3]

5.3 Bias Mitigation

- Models may inadvertently perpetuate biases present in the training data, leading to unfair or inaccurate training content. [4]
- Continuous monitoring and debiasing techniques are necessary to ensure fairness.

5.4 Regulatory Approval

- AI - generated training content may require approval from regulatory bodies, potentially slowing down implementation. [5]
- Demonstrating the reliability and accuracy of AI - generated content to regulators can be challenging.

6. Ethical Considerations

6.1 Transparency

- Clearly communicate the use of AI in generating and delivering training content to all stakeholders.
- Provide mechanisms for employees to understand how personalized recommendations are made.

6.2 Fairness

- Regularly audit the AI system for potential biases in content generation or assessment.
- Ensure that the training system provides equal opportunities for learning and advancement across all employee groups.

6.3 Human Oversight

- Maintain a system of human expert review for critical compliance content.
- Establish clear escalation pathways for complex compliance questions that the AI system cannot adequately address.

6.4 Data Privacy

- Implement robust data protection measures to safeguard employee information used in personalizing training.
- Ensure compliance with relevant data protection regulations (e. g., GDPR) in the collection and use of training data.

7. Future Directions

7.1 Multimodal Learning

- Integrate computer vision techniques to incorporate visual elements (e. g., document layouts, signatures) into compliance training.
- Develop voice - based interfaces for hands - free compliance assistance during operational tasks.

7.2 Federated Learning

- Implement federated learning techniques to allow multiple financial institutions to collaboratively train compliance models without sharing sensitive data.

7.3 Explainable AI (XAI)

- Develop and integrate XAI techniques to provide clear explanations for AI - generated training content and recommendations.
- Create visual tools to help employees understand the reasoning behind compliance decisions.

7.4 Regulatory Technology (RegTech) Integration

- Explore integration with emerging RegTech solutions for real - time compliance monitoring and reporting. [6]
- Develop APIs for seamless data exchange between AI training systems and regulatory reporting platforms.

8. Conclusion

AI - generated regulatory compliance training offers a transformative approach for the banking and payment industry. By leveraging fine - tuned small LLMs, institutions can ensure their employees are well - versed in the latest regulations, enhancing compliance and reducing risk. Despite existing challenges, the potential benefits in terms of efficiency, personalization, and risk mitigation make this an exciting area for future development and implementation.

References

[1] Fernandez, L. (2021). JPMorgan's journey to improving AML efficacy through artificial intelligence. Journal of Financial Compliance, 4 (3), 225 - 237.

[2] Knight, W. (2020). How HSBC is using AI to boost its compliance procedures. MIT Technology Review, 123 (4), 78 - 85.

[3] Financial Action Task Force. (2021). Opportunities and Challenges of New Technologies for AML/CFT. FATF Report.

[4] Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. California Law Review, 104, 671.

[5] European Banking Authority. (2020). Report on Big Data and Advanced Analytics. EBA/REP/2020/01.

[6] Office of the Comptroller of the Currency. (2019). Supporting Responsible Innovation in the Federal Banking System: An OCC Perspective.

[7] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N.,. . . & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

[8] Howard, J., & Ruder, S. (2018). Universal language model fine - tuning for text classification. arXiv preprint arXiv: 1801.06146.

[9] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre - training of deep bidirectional transformers for language understanding. arXiv preprint arXiv: 1810.04805.