The OSCAR PUBLISHING Services
American Journal of Applied Science and Technology

VOLUME Vol.06 Issue 08 2026
PAGE NO. 123-131
DOI 10.37547/ajast/Volume06Issue08-08

Secure-XPAPF: A Cyber-Secure and Explainable Deep Learning Framework for Automated Pronunciation Assessment and Personalized Feedback in Distance Language Learning

Khamdamov Utkir Rakhmatullayevich

Tashkent University of Information Technologies named after Muhammad al-Khorezmi, 108, Amir Temur Av.,100200, Tashkent, Uzbekistan

Turakulov Olim Xolbutayevich

Jizzakh branch of the National University of Uzbekistan named after Mirzo Ulugbek, Jizzakh, Uzbekistan

Abdumalikov Akmaljon Abduxoliq o‘g‘li

Jizzakh branch of the National University of Uzbekistan named after Mirzo Ulugbek, Jizzakh, Uzbekistan

Kayumov Oybek Achilovich

Jizzakh branch of the National University of Uzbekistan named after Mirzo Ulugbek, Jizzakh, Uzbekistan

Akmuradov Baxtiyor Uralovich

Vatan University, 2nd Lane, Shota Rustaveli Street, House 2, Dumbuloq MFY, Yakkasaroy District, Tashkent, Uzbekistan

Received: 24 June 2026; **Accepted:** 20 July 2026; **Published:** 17 August 2026

Abstract: Automated pronunciation assessment has become a core technology for distance foreign language learning, yet existing systems frequently behave as opaque scoring engines and rarely address the cybersecurity risks created by collecting speech, storing learner profiles, generating AI feedback and operating learning analytics at scale. This article proposes Secure-XPAPF, a cyber-secure explainable deep learning framework for automated pronunciation assessment and personalized corrective feedback. The framework integrates automatic speech recognition, pronunciation error detection, pronunciation quality scoring, explainable AI, personalized feedback recommendation, learning analytics, longitudinal progress modeling and a cross-cutting cybersecurity layer. Technically, Secure-XPAPF combines self-supervised speech representations, Transformer-based multi-task pronunciation modeling, phoneme-level error diagnosis, SHAP-style acoustic attribution, attention and saliency visualization, evidence-bound large-language-model feedback, privacy-preserving learning analytics, differential-privacy-aware federated training, adversarial audio robustness testing and zero-trust access governance. Pedagogically, it transforms numerical scores into learner-facing explanations, targeted corrective exercises and longitudinal dashboards for teachers and learners. From a cybersecurity perspective, the framework maps the full attack surface of distance pronunciation learning, including speech data leakage, metadata re-identification, model inversion, membership inference, data poisoning, adversarial audio, prompt injection, insecure LLM output handling and analytics dashboard abuse. The evaluation design combines regression, classification, explanation fidelity, recommendation ranking, learning gain, trust, privacy leakage and security resilience metrics. Illustrative, simulation-calibrated results are provided only as a reproducible reporting template and indicate how a deployed system could compare against GOP, CNN-BiLSTM, Whisper-based and wav2vec2-based baselines. The contribution

is a unified, auditable and security-by-design framework that connects speech AI, explainable pedagogy and cyber-resilient learning analytics for high-stakes distance language learning environments.

Keywords: Cyber-secure pronunciation assessment; explainable AI; automatic speech recognition; pronunciation error detection; personalized corrective feedback; learning analytics; differential privacy; federated learning; LLM security; distance foreign language learning .

INTRODUCTION:

Distance foreign language learning has expanded the need for pronunciation assessment systems that can evaluate learner speech continuously, consistently and at scale. In conventional classrooms, teachers can observe articulation, listen to repeated attempts and provide contextual feedback. In online environments, this interaction is often mediated by digital platforms that must process speech recordings, infer pronunciation quality, recommend practice and maintain learner progress histories. The technical promise of automated pronunciation assessment is therefore substantial, but the educational and cyber-risk profile is also higher because speech is biometric-like, culturally revealing and difficult to anonymize completely. Recent advances in self-supervised speech representation learning have changed the technical foundations of pronunciation assessment. wav2vec 2.0, HuBERT, WavLM, Whisper and related models learn high-dimensional acoustic representations that can be adapted for alignment, error diagnosis and scoring. At the same time, Transformer-based pronunciation models such as GOPT have shown that multi-aspect and multi-granularity scoring is more suitable than a single opaque score. However, a high correlation with human raters does not automatically make a system pedagogically useful, explainable, secure or privacy-preserving. The cybersecurity dimension is no longer optional. A pronunciation learning platform collects raw audio, device metadata, learner identity attributes, feedback histories, teacher comments, dashboard events and potentially LLM-generated recommendations. These assets can be attacked through classical application vulnerabilities, insecure APIs, credential misuse, model extraction, membership inference, data poisoning, adversarial audio samples, prompt injection and unauthorized dashboard access. If the platform is used by minors, international students or high-stakes language learners, the social and legal consequences of leakage or misclassification become even more serious.

Table 1. Acronyms and functional interpretation used throughout the manuscript.

Acronym	Meaning	Role in Secure-XPAPF
APA	Automatic Pronunciation Assessment	Estimates utterance-, word- and phoneme-level pronunciation quality.
ASR	Automatic Speech Recognition	Aligns spoken input with expected text and phoneme sequences.
PED	Pronunciation Error Detection	Detects substitutions, deletions, insertions, stress errors and prosodic deviations.
PQS	Pronunciation Quality Scoring	Transforms acoustic-phonetic evidence into calibrated scores.
XAI	Explainable Artificial Intelligence	Explains why a score or error label was assigned.
PFR	Personalized Feedback Recommendation	Maps detected weaknesses to adaptive corrective tasks.
LA	Learning Analytics	Tracks learner behavior, progress and instructional risk.
LLPM	Longitudinal Learning Progress Modeling	Models pronunciation mastery over repeated practice.
DP-FL	Differentially Private Federated Learning	Improves speech models while limiting raw data centralization and privacy leakage.
CSL	Cybersecurity Layer	Provides confidentiality, integrity, availability, privacy, accountability and resilience controls.

Secure-XPAPF is proposed in response to this convergence. It treats pronunciation assessment as an instructional decision pipeline that must be accurate, interpretable, personalized, longitudinal and cyber-resilient. The framework is security-by-design because it embeds data minimization, encryption, role-based access, privacy-preserving model training, audit logging, LLM guardrails and adversarial robustness testing into each technical layer rather than adding security as a final compliance checklist.

Research problem

The central research problem is the absence of an integrated framework that simultaneously supports explainable pronunciation assessment, personalized corrective feedback, longitudinal learning analytics and cybersecurity governance. Existing systems often optimize one part of the pipeline: ASR alignment, phoneme-level error detection, pronunciation scoring, feedback generation or analytics visualization. In practice, however, distance learners and teachers require a complete chain from secure speech capture to trustworthy feedback and measurable progress.

This fragmentation creates five limitations. First, score-only APA systems do not identify the acoustic or phonetic causes of weak performance. Second, error detection outputs are often not translated into adaptive practice. Third, explainability methods are rarely converted into learner-facing language. Fourth, learning analytics dashboards may collect sensitive behavioral traces without sufficient privacy controls. Fifth, LLM-based feedback can introduce new attack surfaces, including prompt injection, hallucinated correction, sensitive information disclosure and insecure output handling.

Research gaps

- Gap 1 - Scoring without explanation: APA models increasingly produce numerical scores but rarely provide local acoustic-phonetic evidence that teachers and learners can inspect.
- Gap 2 - Diagnosis without personalization: PED models can detect phoneme errors, but many systems do not transform those detections into individualized learning paths.
- Gap 3 - Feedback without cybersecurity: LLM-enabled feedback is promising, but evidence-bound generation, prompt-injection resistance and sensitive-data protection remain underdeveloped.
- Gap 4 - Analytics without privacy engineering: Longitudinal learning analytics can improve instruction, yet they introduce re-identification, profiling and unauthorized access risks.
- Gap 5 - Accuracy without resilience: APA research commonly reports MAE, RMSE, PCC, accuracy or F1 but rarely reports adversarial audio robustness, privacy leakage, model extraction resistance or data poisoning sensitivity.
- Gap 6 - Siloed evaluation: Predictive, pedagogical, explanatory, recommendation and cybersecurity evaluations are seldom integrated in one protocol.

Research questions

1. How can self-supervised ASR and speech representations be used to assess pronunciation quality at utterance, word and phoneme levels?
2. How can pronunciation model decisions be explained through acoustic-phonetic evidence that is meaningful to learners and teachers?
3. How can detected pronunciation errors be transformed into personalized corrective feedback and adaptive exercises?
4. How can longitudinal learning analytics model pronunciation development over repeated distance-learning sessions?
5. How can cybersecurity, privacy and LLM-safety controls be embedded into the entire pronunciation assessment pipeline?
6. How should Secure-XPAPF be evaluated across predictive accuracy, diagnostic reliability, explainability, recommendation quality, learning gain, privacy leakage and security resilience?

Contributions

- C1: A cyber-secure eight-layer Secure-XPAPF architecture that unifies ASR, PED, PQS, XAI, PFR, LA, LLPM and cybersecurity governance.

- C2: A multi-task pronunciation model that jointly estimates pronunciation quality and phoneme-level diagnostic labels while supporting calibrated uncertainty.
- C3: A learner-facing explainability mechanism that combines SHAP-style feature attribution, attention visualization, acoustic saliency and explanation fidelity checks.
- C4: A hybrid feedback engine that combines deterministic pedagogical rules with evidence-bound LLM generation constrained by safety, privacy and hallucination-control policies.
- C5: A privacy-preserving learning analytics model that tracks weak phonemes, growth curves and mastery states while minimizing unnecessary learner data exposure.
- C6: A cybersecurity threat model for pronunciation learning platforms covering speech data leakage, metadata linkage, model inversion, membership inference, poisoning, adversarial audio, prompt injection and dashboard abuse.
- C7: A multidimensional evaluation protocol including accuracy, diagnostic F1, explanation fidelity, NDCG, learning gain, trust score, privacy leakage index and security resilience score.
- C8: A reproducible manuscript blueprint with formulas, tables and protocol-ready results templates suitable for future Q1-level empirical submission once real data are available.

Table 2. Research problem mapped to Secure-XPAPF design responses.

Limitation in existing systems	Secure-XPAPF response	Primary evaluation indicator
Opaque pronunciation score	Local acoustic attribution and phoneme-level explanations	Explanation fidelity, trust score
Error diagnosis not linked to practice	Personalized corrective feedback and adaptive exercises	NDCG, MAP, learning gain
No longitudinal modeling	Dynamic mastery state for each phoneme and prosodic skill	Learning curve slope, retention gain
Limited privacy governance	Data minimization, DP-FL, encryption, access control	Privacy leakage index, audit completeness
Vulnerable LLM feedback	Prompt firewall, evidence grounding, output validation	Injection success rate, unsafe output rate
No adversarial robustness reporting	Audio perturbation testing and anomaly detection	Robust F1, adversarial degradation rate

LITERATURE REVIEW

Automatic pronunciation assessment has developed from likelihood-based Goodness of Pronunciation methods toward neural and self-supervised speech models. Traditional GOP compares acoustic likelihoods between canonical and competing phoneme hypotheses; it is interpretable at the phoneme level but depends heavily on alignment quality and ASR acoustic modeling assumptions. Deep learning approaches improve representation capacity by using CNN, LSTM, BiLSTM, CTC, attention and Transformer architectures, yet their internal decision logic is harder to interpret.

The Speechocean762 corpus has become an important public benchmark because it provides non-native English recordings with sentence-, word- and phoneme-level annotations. The GOPT model demonstrated the importance of multi-aspect and multi-granularity modeling by jointly addressing accuracy, fluency, prosody and completeness. SSL-based models further improved the field by using learned representations from wav2vec 2.0, HuBERT and related architectures, which can encode pronunciation-relevant information beyond handcrafted MFCC and Mel features. Despite these advances, APA remains educationally incomplete when it returns only a numerical score. Learners need to know whether their low score is caused by

segmental substitution, vowel length, stress displacement, syllable deletion, intonation mismatch or fluency instability. Secure-XPAPF therefore treats APA as a diagnostic and instructional problem rather than as score regression alone.

Pronunciation error detection and phoneme-level diagnosis

Pronunciation error detection identifies deviations between canonical phoneme sequences and learner realizations. Typical labels include correct realization, substitution, deletion, insertion, distorted articulation, stress error and prosodic deviation. Segment-level diagnosis is essential because a learner who consistently confuses /theta/ and /s/ requires a different intervention than a learner who has low fluency but stable segmental articulation. Recent models use forced alignment, CTC, encoder-decoder attention, phone posterior features, anti-phone modeling, graph-based articulation features and multi-view fusion. Secure-XPAPF extends this line of work by linking PED outputs to explainability, feedback and longitudinal analytics. Each detected error is stored not merely as a label but as an evidence object containing the target phoneme, acoustic segment, confidence, explanation vector, recommended exercise type and historical frequency.

Explainable AI in education and speech systems

Explainable AI is necessary in educational assessment because decisions influence learner motivation, teacher intervention and institutional trust. General XAI methods such as LIME and SHAP approximate model behavior through local perturbations or additive feature contributions, while attention visualization and acoustic saliency can highlight temporal regions that drive speech-model decisions. Secure-XPAPF distinguishes technical explanations from instructional explanations. Technical explanations support auditability, debugging and model comparison; instructional explanations support learner action. A local explanation must be faithful to the model, stable across small benign perturbations and understandable to the learner. A global explanation must reveal which features, phonemes and learner groups drive system behavior across the dataset.

Personalized corrective feedback and intelligent tutoring

Personalized learning systems adapt content, sequence, difficulty and feedback according to learner state. In pronunciation instruction, personalization should consider L1 background, current proficiency, weak phoneme profile, practice history, response to previous feedback, affective state and teacher priorities. A one-size-fits-all correction is inefficient because different learners may produce the same target sentence with different acoustic causes of error. LLM-based feedback can improve naturalness and pedagogical detail, but it also creates cybersecurity and safety risks. A generic LLM may hallucinate phonetic explanations, expose hidden system prompts, leak learner data, produce culturally insensitive feedback or obey malicious prompt injection embedded in user text. Secure-XPAPF therefore uses evidence-bound generation: the LLM receives only structured diagnostic facts, permitted feedback templates and safety constraints, and its output is validated before display.

Learning analytics transforms isolated assessment events into developmental trajectories. In pronunciation learning, analytics can identify persistent weak phonemes, changes in fluency, practice-response patterns, exercise completion, time-to-mastery and intervention effects. However, analytics also expands the data surface: a system may store repeated recordings, timestamps, device identifiers, location-like metadata, behavioral logs and teacher notes.

Privacy and data protection must therefore be implemented throughout the analytics lifecycle. Secure-XPAPF follows data minimization, purpose limitation, role separation, consent logging and retention control. Where institutional deployment permits, federated learning and differential privacy can reduce raw data centralization and limit inference from model updates.

The framework differs from conventional APA pipelines in four ways. First, it treats security and privacy as modeling constraints rather than administrative add-ons. Second, it treats explanations as pedagogical objects, not only technical artifacts. Third, it links each detected error to an adaptive feedback pathway. Fourth, it stores progress signals as privacy-minimized longitudinal states instead of raw lifelong surveillance traces.

Table 3. Secure-XPAPF layers and embedded cybersecurity controls.

Layer	Pedagogical/AI function	Security-by-design control
1. Secure speech acquisition	Captures learner audio and task context.	Consent, TLS, device integrity, metadata minimization.
2. Feature extraction	Computes MFCC, Mel spectrogram and SSL embeddings.	On-device feature extraction where feasible; raw-audio minimization.
3. ASR and alignment	Aligns audio with text and phoneme targets.	Input validation, anomaly detection, robust alignment confidence.
4. Pronunciation quality scoring	Predicts multi-level scores.	Uncertainty calibration and model integrity verification.
5. Pronunciation error detection	Classifies phoneme and prosody errors.	Poisoning-resistant training and adversarial testing.
6. Explainability	Generates local and global explanations.	Explanation audit, stability check, no sensitive feature leakage.
7. Personalized feedback	Recommends corrective practice and explanations.	Evidence-bound LLM prompts, prompt firewall, output validation.
8. Learning analytics	Tracks progress and mastery.	Pseudonymization, retention limits, role-based dashboard access.

Dataset Construction and Ethical-Cybersecurity Governance

A publishable Secure-XPAPF evaluation should combine a public benchmark such as Speechocean762 with an institutionally collected distance-learning corpus. The institutional corpus should be constructed only after ethics approval, informed consent and a data protection impact assessment. The recommended corpus size is 300 to 500 learners, 8 to 10 weekly sessions and a balanced set of read-aloud sentences, minimal pairs, stress tasks and spontaneous short responses.

The corpus design must separate research identifiers from direct identity data. Raw recordings should be encrypted at rest, transmitted over TLS, stored with pseudonymous learner IDs and retained only for approved purposes. Audio features and model embeddings should be treated as sensitive because they may still contain speaker-identifying information. Teachers should access only the learner information required for instruction, while researchers should access de-identified data whenever possible. Annotation should be performed by at least three language specialists. Each utterance should receive sentence-level, word-level and phoneme-level ratings for accuracy, fluency, completeness and prosody. Inter-annotator agreement should be reported using Cohen's kappa for pairwise labels, Fleiss' kappa for multi-rater categorical labels and intraclass correlation for continuous scores. Disagreements should be adjudicated through a documented protocol.

Explainable AI Module

The XAI module generates local and global explanations. A local explanation answers why the model produced a score for a particular learner utterance. A global explanation describes which acoustic features, phoneme types, learner proficiency levels or task categories influence the model across the dataset. In Secure-XPAPF, explanation generation is constrained by fidelity, stability and privacy requirements. Feature attribution is computed over acoustic descriptors, SSL embedding dimensions, temporal segments and aligned phoneme units.

Because raw SSL dimensions are not directly meaningful to learners, the system maps them to interpretable evidence categories such as duration mismatch, spectral shift, voicing difference, stress deviation, pause pattern and intonation contour.

Equation (6). Additive local explanation:

$$g_i(\mathbf{x}) = \phi_{0,i} + \sum_{m=1}^{M} \phi_{m,i}\, q_m(\mathbf{x}), \qquad \sum_{m=1}^{M} |\phi_{m,i}| = 1$$

Here phi_m is the contribution of interpretable evidence q_m to the local pronunciation decision. The learner-facing explanation is generated only from approved evidence categories.

Equation (7). Explanation fidelity:

$$\mathrm{Fid}_i = 1 - \frac{|f(\mathbf{x}_i) - g_i(\mathbf{x}_i)|}{|f(\mathbf{x}_i)| + \varepsilon}$$

High fidelity indicates that the explanation approximates the underlying model decision locally.

Equation (8). Explanation stability:

$$\mathrm{Stab}_i = 1 - \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(g_i(\mathbf{x}_i + \delta_b) - \overline{g}_i\right)^2}$$

A robust explanation should not change substantially under small benign perturbations of the audio signal.

Table 4. Explanation objects generated for each learner attempt.

Explanation object	Technical form	Learner-facing transformation
Acoustic saliency	Time-frequency saliency map	Highlights the segment where the pronunciation diverged.
Phoneme attribution	Contribution score per aligned phoneme	Shows which sound lowered the score.
Feature importance	Duration, F0, energy, spectral and voicing cues	Explains whether the issue was stress, length, voicing or articulation.
Counterfactual hint	Minimal change needed to improve score	Suggests a precise corrective action.
Global pattern	Aggregated weak phoneme ranking	Helps teachers design group-level intervention.

Personalized Corrective Feedback Engine

The feedback engine converts diagnostic evidence into corrective instruction. Secure-XPAPF uses a hybrid architecture: deterministic rules provide phonetic correctness and curriculum alignment, while a constrained LLM converts structured evidence into natural language explanations and motivational guidance. The LLM is never allowed to invent acoustic evidence; it receives a structured evidence packet and approved pedagogical templates. Personalization depends on the learner state vector, which includes current weak phonemes, recent improvement slope, exercise completion, L1-specific difficulty profile, confidence trend and teacher-defined objectives. The system recommends exercises that maximize expected learning gain while controlling cognitive load and avoiding excessive repetition.

CONCLUSION

This article proposed Secure-XPAPF, a cyber-secure explainable deep learning framework for automated pronunciation assessment and personalized corrective feedback in distance foreign language learning. The framework extends conventional APA by integrating ASR, PED, PQS, XAI, PFR, LA, LLPM and cybersecurity governance into a single auditable architecture. Its main contribution is not merely a new neural architecture but a complete system logic that connects score prediction, diagnostic explanation, adaptive pedagogy, privacy-preserving analytics and cyber-resilience.

- Secure-XPAPF provides an eight-layer architecture for secure speech acquisition, representation learning, scoring, diagnosis, explanation, feedback, analytics and cybersecurity governance.
- The framework introduces a multi-objective modeling formulation that combines scoring accuracy, error detection, explanation

regularization, privacy protection and adversarial robustness.

- The cybersecurity layer addresses speech data leakage, metadata linkage, membership inference, model inversion, data poisoning, adversarial audio, prompt injection and dashboard abuse.
- The personalized feedback engine uses evidence-bound LLM generation, ensuring that learner-facing explanations are grounded in detected acoustic-phonetic evidence.
- The evaluation protocol extends Q1-style APA reporting by adding explanation, recommendation, learning gain, privacy leakage and security resilience metrics.
- The illustrative results tables provide a transparent reporting template and must be replaced by real experimental findings before journal submission.

REFERENCES

1. Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. Proceedings of the ACM SIGSAC Conference on Computer and Communications Security.
2. Arrieta, A. B., Diaz-Rodriguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82-115.
3. Baevski, A., Zhou, Y., Mohamed, A., & Auli, M. (2020). wav2vec 2.0: A framework for self-supervised learning of speech representations. Advances in Neural Information Processing Systems.
4. Cerratto Pargman, T., & McGrath, C. (2021). Mapping the ethics of learning analytics in higher education: A systematic literature review of empirical research. Journal of Learning Analytics, 8(2), 123-139.
5. Feng, T., Peri, R., & Narayanan, S. (2022). User-level differential privacy against attribute inference attack of speech emotion recognition in federated learning. arXiv:2204.02500.
6. El Kheir, Y., & Glass, J. (2023). Automatic pronunciation assessment - A review. Findings of the Association for Computational Linguistics: EMNLP.
7. Fredrikson, M., Jha, S., & Ristenpart, T. (2015). Model inversion attacks that exploit confidence information and basic countermeasures. Proceedings of ACM CCS.
8. Gong, Y., Chen, Z., Chu, I. H., Chang, P., & Glass, J. (2022). Transformer-based multi-aspect multi-granularity non-native English speaker pronunciation assessment. arXiv:2205.03432.
9. Han, H., et al. (2026). Multi-granularity Interactive Attention Framework for automatic pronunciation assessment. Proceedings of AAAI Conference on Artificial Intelligence.
10. Karimov, A., Baker, R. S., et al. (2024). Ethical considerations and student perceptions of learning analytics data collection and use. HICSS.
11. McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Aguera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. AISTATS.
12. Mothukuri, V., Khare, P., Parizi, R. M., Pouriyeh, S., Dehghantanha, A., & Srivastava, G. (2021). Federated-learning-based anomaly detection for IoT security attacks. IEEE Internet of Things Journal.
13. National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1.
14. National Institute of Standards and Technology. (2024). The NIST Cybersecurity Framework (CSF) 2.0. NIST Cybersecurity White Paper.
15. Qin, Y., Carlini, N., Cottrell, G., Goodfellow, I., & Raffel, C. (2019). Imperceptible, robust, and targeted adversarial examples for automatic speech recognition. International Conference on Machine Learning.
16. Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2023). Robust speech recognition via large-scale weak supervision. International Conference on Machine Learning.
17. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. IEEE ICCV.

18. Shoemate, M., Jett, K., Cowan, E., Colbath, S., Honaker, J., & Muthukumar, P. (2022). Sotto Voce: Federated speech recognition with differential privacy guarantees. arXiv:2207.07816.

19. Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). Membership inference attacks against machine learning models. IEEE Symposium on Security and Privacy.

20. Tramer, F., Zhang, F., Juels, A., Reiter, M. K., & Ristenpart, T. (2016). Stealing machine learning models via prediction APIs. USENIX Security Symposium.

21. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems.