CS + Music Research Projects for Advanced High School Students

An advanced high school student with machine learning experience and piano expertise has exceptional opportunities to contribute meaningfully to the rapidly evolving intersection of computer science and music. Recent developments in AI music generation, improved educational resources, and accessible pre-trained models have made sophisticated music technology projects achievable within a school year timeframe while maintaining genuine research value.

Current Research Landscape: Why Now is the Perfect Time

The field of Music Information Retrieval (MIR) and AI music generation has reached a remarkable inflection point for student researchers. (PubMed Central +2) **Meta's MusicGen (2023), Google's enhanced Magenta platform,** (arXiv) **and the release of massive datasets like the 666,000-song Chordonomicon** (arXiv) have democratized access to cutting-edge music AI capabilities. (withgoogle +2) University labs at NYU MARL, Stanford CCRMA, and Georgia Tech actively recruit undergraduate researchers, (NYU) (NYU Steinhardt) while the ISMIR (International Society for Music Information Retrieval) community explicitly welcomes student contributions through dedicated educational tracks. (Wikipedia +4)

Three factors make this field particularly suitable for talented high school students. First, the intersection rewards both technical and musical expertise—your piano background becomes a genuine research advantage rather than merely helpful context. Second, clear evaluation metrics exist for most music tasks, from note-level precision in transcription to harmonic accuracy in generation. (arXiv) Third, abundant open-source tools and datasets eliminate traditional barriers to entry.

Current breakthrough areas include transformer-based music generation, real-time performance analysis, and accessibility-focused music technology. (MDPI) Recent papers demonstrate that high-quality results are achievable with consumer hardware and publicly available datasets, particularly for piano-focused applications where your musical expertise provides crucial validation capabilities. (Nature)

Essential Toolkit: Frameworks and Libraries

Core Music + ML Platforms

Google Magenta serves as the cornerstone platform, offering both educational entry points and research-grade capabilities. Magenta Studio provides immediate experimentation opportunities, while the underlying TensorFlow models enable custom development. (withgoogle +6) **The platform's strength lies in real-time MIDI integration**—essential for piano-based projects—and comprehensive documentation including Colab notebooks for immediate experimentation.

Music21 (MIT) excels at symbolic music analysis and manipulation, perfectly suited for projects requiring music theory integration. (Librosa +3) Its extensive corpus of classical works and sophisticated harmonic analysis capabilities make it ideal for piano repertoire studies. (Music21) The library's object-oriented design maps naturally to musical concepts, reducing the learning curve for musician-programmers. (music21 +3)

PyTorch and TensorFlow provide the foundation for custom ML models, with extensive music-specific examples available. **Hugging Face's growing collection of music models, particularly MusicGen, offers**

state-of-the-art generation capabilities through simple APIs, arXiv enabling rapid prototyping and fine-tuning for specific applications. Hugging Face TensorFlow

Audio Processing and Real-Time Systems

Librosa leads audio analysis capabilities, offering everything from basic feature extraction to advanced spectral analysis. Medium +2 Its integration with the scientific Python stack and comprehensive visualization tools make it essential for any audio-based project. Analytics Vidhya +2 For real-time applications, **PyAudio and the Web Audio API enable low-latency processing suitable for live performance systems**. ui.dev +2

The combination of **pretty_midi for MIDI manipulation, Tone.js for web-based audio, and various synthesis libraries** creates a complete development environment. ui.dev +2 Hardware integration through Arduino or Raspberry Pi connects to physical controllers and sensors, expanding project possibilities beyond software-only solutions.

High-Impact Project Categories

Music Generation and AI Composition

Intelligent Piano Accompaniment Generator represents an ideal fusion of technical challenge and musical application. This project leverages conditional VAEs or Transformer architectures to generate piano accompaniment for given melodies, with implementation spanning 6-8 months. PubMed Central The technical approach involves collecting classical piano scores, separating melody and accompaniment tracks using music21, and training models to learn harmonic relationships. music21 Librosa

Your piano expertise becomes crucial for curating training data and evaluating results—a task where musical knowledge provides genuine advantage over purely technical approaches. **Success metrics include harmonic accuracy, rhythmic coherence, and style consistency**, with the potential for real-time demonstration during competitions.

Bach-Style Invention Generator offers more advanced opportunities, implementing counterpoint constraints within neural architectures. Magenta TensorFlow This project combines transformer models with music theory rules, creating two-voice piano inventions while maintaining voice independence. PubMed Central Magenta The 8-10 month timeline includes comprehensive analysis tools using music21, making it suitable for both technical and musicological evaluation. music21 Librosa

Real-Time Piano Learning Companion addresses practical educational needs through AI-powered practice assistance. The system combines audio transcription, contextual accompaniment generation, and performance analysis, with implementation requiring audio processing, real-time systems expertise, and user interface design. Magenta Hugging Face This project offers strong competition potential through live demonstration capabilities.

Music Analysis and Performance Enhancement

Automatic Music Transcription (AMT) projects leverage your piano expertise for specialized transcription systems. Using Google's pre-trained Onsets and Frames model as a foundation, projects can focus on classical piano repertoire with enhanced accuracy through domain-specific training.

MIDI.org TensorFlow **The MAESTRO dataset provides 200+ hours of piano performances with aligned MIDI/audio**, offering comprehensive training data. Papers with Code TensorFlow

Technical requirements include TensorFlow/Magenta familiarity, audio processing with Librosa, and MIDI manipulation skills. Medium +3 **Current state-of-the-art achieves ~75% accuracy on standard datasets**, GitHub with opportunities for improvement through piano-specific optimization and musical knowledge integration. GitHub

Piano Performance Analysis projects examine expressive performance characteristics using MIDI data from professional recordings. Nature These systems analyze timing variations, dynamic patterns, and articulation styles, comparing student performances against professional benchmarks. Nature The **ATEPP dataset provides 11,677 performances by virtuoso pianists**, enabling comprehensive style analysis and practice feedback systems. Papers with Code

Success metrics focus on correlation with expert ratings and classification accuracy for different performance styles. Projects can extend to real-time feedback systems, progress tracking analytics, and comparative analysis tools for different interpretations of classical works.

Music Information Retrieval and Classification

Genre Classification projects provide excellent entry points with clear success metrics. Interview Query GeeksforGeeks Classical music period classification (Baroque, Classical, Romantic, Modern) leverages your historical knowledge while developing CNN architectures for spectrogram analysis. Interview Query GeeksforGeeks **The GTZAN dataset offers standardized evaluation, though piano-specific datasets often yield more meaningful results for classical applications**. ResearchGate +3

Chord Recognition and Harmonic Analysis projects address complex jazz harmonies and classical modulations, areas where traditional systems struggle. Technical approaches combine chroma feature extraction with deep learning models, enhanced by music theory knowledge for post-processing. LinkedIn interviewquery Your piano background enables evaluation of complex harmonic progressions that challenge purely computational approaches.

Music Similarity and Recommendation systems can focus on classical repertoire, incorporating composer styles, technical difficulty, and musical structure. Interview Query These projects combine content-based filtering using audio features with collaborative approaches, creating sophisticated recommendation engines for piano students and teachers. interviewquery Interview Query

Interactive Applications and Real-Time Systems

Web-Based Music Applications using Tone.js and Web Audio API create impressive demonstrations while developing practical skills. github +5 An **Adaptive Piano Practice Companion** provides real-time performance feedback through browser-based audio analysis, combining machine learning with interactive visualization. Magenta

AR Piano Learning Environment projects integrate computer vision with music education, overlaying visual guidance on physical or virtual pianos. Implementation uses ARCore/ARKit with hand tracking, creating compelling demonstration opportunities through smartphone or tablet interfaces.

Gesture-Controlled MIDI Systems combine hardware integration with musical expression, using sensors or computer vision to control musical parameters. Instructables These projects offer impressive visual appeal for competitions while exploring innovative human-computer interaction paradigms.

Project Selection Framework

Technical Feasibility Assessment

Beginner projects (2-3 months) include music genre classification, basic MIDI analysis, and simple beat tracking systems. (LinkedIn +2) These build fundamental skills while producing meaningful results suitable for initial portfolio development. (GeeksforGeeks) (Polygence)

Intermediate projects (4-6 months) encompass piano practice assistants, automatic transcription systems, and music recommendation engines. (Medium +2) These require more sophisticated ML implementation but remain achievable with dedicated effort and appropriate guidance.

Advanced projects (6-12 months) involve expressive performance synthesis, multi-track analysis systems, and AI-assisted composition tools. These approach research-level complexity while leveraging your piano expertise for unique contributions.

Evaluation and Success Metrics

Objective metrics provide quantifiable assessment through established benchmarks. Note-level precision/recall for transcription, classification accuracy for genre recognition, and harmonic accuracy for generation tasks offer clear success indicators comparable to published research. (GitHub) (arXiv)

Subjective evaluation leverages your musical knowledge through playability assessment, musical expression evaluation, and style consistency analysis. Human evaluation frameworks include blind listening tests, expert panel reviews, and comparative analysis with existing compositions. (arXiv +2)

Research Impact and Applications

Educational technology applications address real needs in music instruction through intelligent tutoring systems, automated assessment tools, and personalized practice recommendations. (Nature +2) These projects often generate significant interest from music educators and technology integrators. (National Association for Music...)

Music industry relevance includes content-based discovery systems, automatic metadata generation, and performance rights identification. Projects with commercial applications demonstrate practical value beyond academic interest.

Accessibility and inclusion opportunities focus on tools for hearing-impaired musicians, simplified interfaces for non-technical users, and cross-cultural music understanding systems. These areas offer meaningful social impact alongside technical achievement. (ScienceDirect)

Implementation Strategy and Timeline

Phase 1: Foundation Building (Months 1-2)

Establish development environment with Python, essential libraries, and MIDI interface capabilities. (Librosa +2) Download and explore key datasets including MAESTRO, GTZAN, and relevant classical music collections. (TensorFlow) (tensorflow) Create data preprocessing pipelines and establish training/validation protocols.

Phase 2: Model Development (Months 3-6)

Begin with simple implementations using pre-trained models, then progressively increase complexity. Regular consultation with piano teachers or advanced musicians ensures musical validity of technical approaches. Focus on rapid iteration and clear evaluation metrics.

Phase 3: Advanced Features (Months 7-10)

Implement real-time capabilities, style transfer functionality, and interactive elements based on initial results. (Magenta) (Hugging Face) Optimize performance for live demonstration and add user control mechanisms. Develop comprehensive evaluation frameworks combining technical and musical assessment.

Phase 4: Presentation and Documentation (Months 11-12)

Prepare competition materials with emphasis on live demonstration capabilities. Create comprehensive documentation explaining technical methods, musical insights, and broader implications. Develop presentation materials suitable for both technical and general audiences.

Competition and Recognition Opportunities

Science fairs at regional and national levels increasingly recognize interdisciplinary projects combining technology with creative domains. (Science Buddies) (Inspirit AI) **Music technology competitions** and **GRAMMY Camp submissions** specifically target innovative music applications. (Polygence)

Academic conferences including ISMIR offer student research presentations and late-breaking demo sessions. (ISMIR 2023 +3) **Open source contributions** to established projects like Magenta demonstrate technical capability while building community connections. (GitHub) (MarkTechPost)

University research opportunities often begin with high school projects, particularly when demonstrating genuine interdisciplinary expertise. (Hamilton College) Many programs specifically recruit students with both technical and artistic backgrounds for vertically integrated projects.

Success Factors and Common Pitfalls

Success requires balancing technical sophistication with musical validity. Projects that ignore musical principles in favor of purely technical metrics often produce impressive numbers but musically meaningless results. (PubMed Central) (arXiv) Conversely, musically interesting ideas require sufficient technical rigor for competitive evaluation.

Start with achievable goals and expand systematically. The most successful student projects begin with simple, working implementations and add complexity incrementally. **Perfect musical outputs matter less than demonstrating systematic improvement and clear evaluation frameworks.**

Leverage your unique combination of skills. The intersection of piano expertise and ML capability provides genuine research advantages that purely technical approaches cannot match. Focus on problems where musical knowledge enables better solutions rather than competing on purely computational grounds. (Hamilton College)

Your piano background and ML experience position you perfectly to contribute meaningfully to this exciting research area. The combination of accessible tools, clear evaluation metrics, and genuine research opportunities creates an ideal environment for ambitious interdisciplinary projects with both academic merit and practical impact. Magenta +2