This study incorporates an advanced state space model tailored for audio signal analysis. This model architecture draws upon the state space framework. State space layers are the main component of the model, which capture the dynamic behavior of audio signals over time. These layers are parameterized by matrices A, B, C, and D, which are adapted throughout the training process. A: Governs the state evolution within a layer, B: Modulates the influence of the input signal on the state, C: Transforms the state into an output signal for subsequent layers and finally D: Provides a direct pathway for the input signal to influence the output. The model features a stack of state space layers, precisely four in this extended configuration, to ensure a comprehensive analysis of the input audio features. The audio data, once pre-processed and structured into tensors, is fed through the state space layers in sequence. Each layer updates its state based on both the current input and its previous state, ensuring that temporal dependencies are maintained and leveraged. Concluding the sequence of state space layers is a linear transformation module that maps the last layer's hidden features to the final output. This output is configured to align with various analysis objectives, from identifying key audio characteristics to categorizing different sounds.