This thesis deals with multistage stochastic programs and explores the dependence of the obtained objective value on the chosen structure of the scenario tree. In particular, the scenario trees are built using the moment matching method, a multistage mean-CVaR model is formulated and a reinforcement learning agent is trained on a set of historical financial data to choose the best scenario tree structure for the mean-CVaR model. For this purpose, we implemented a custom reinforcement learning environment. Further an inclusion of a penalty term in the reward obtained by the agent is proposed to avoid scenario trees that are too complex. The reinforcement learning agent is then evaluated against an agent that chooses the scenario tree structure at random and outperforms the random agent. Further the structure of scenario trees chosen by the reinforcement learning agent is analyzed.