Neural networks have already proven to be able to cope with noisy and unstructured data like hand-written texts, images, sounds, classification of real-world objects based on the incomplete description, and many others. A few endeavours have been made to use Neural Networks for the heuristic learning task. In the standard setting, a course of action of highlights is figured for each state in the preparation set just as the ideal separation to-go to the closest physical state and the system is then used to take in mapping from highlights to remove the gauge. After the learning methodology is done, the framework is used as a heuristic separation estimator together with an informed forward search algorithm like A^\* or IDA\*. This project aims to use heuristic learning to construct an automatically solve a given instance of Rubik’s cube problem. Heuristics that are found out along these lines provide no guarantees on admissibility. The objective of learning is with the goal that the heuristic would be close to the real value but not necessarily always admissible — for example, smaller than the real distance-to-go. Since the search algorithm with an inadmissible heuristic does not typically guarantee to find optimal solutions, this approach is just reasonable in situations where close-to-optimal solutions are sufficient. It is, in any case, conceivable to ensure optimality even with a prohibited heuristic by adjusting the inquiry strategy. Used inadmissible values for larger scramble values; these actions were successful in solving up to 25 moves of scrambles. Furthermore, unacceptable heuristics did not significantly affect the optimal solution. To test this, checked whether the returned length of the solution exceeds the number of scramble moves, the solutions backed have always been optimum, while admissible compromise the optimality of the solution. The learned heuristics were generally perfect compared to those produced by humans. Note that the training process with large amounts of resources can be performed once in this proposed context, and the former model can then be applied by a low-end machine to solve problems effectively.