**Solving Sudoku puzzles with Convolutional Neural Network**Grégory Stehlík Buffard

1. **Abstract**

In this study, we explore the feasibility of employing a deep convolutional neural network (CNN) to accurately solve incomplete Sudoku puzzles. Our approach involved the development of a three-layer CNN, trained on a comprehensive dataset comprising 9 million Sudoku puzzles, each including both unsolved and solved states. Given the constraints imposed by hardware limitations, our training strategy was iterative, utilizing subsets of 10,000 puzzles sequentially to mitigate the risk of overfitting. This methodical process allowed for the incremental training of the network across the entirety of the 9 million puzzle dataset. The performance of our model, as measured by its final accuracy, demonstrates the effective resolution of the posed problem, achieving a success rate deemed satisfactory against predefined criteria. This outcome not only confirms the viability of deep learning models in solving complex logical puzzles but also highlights the potential for similar methodologies to be applied in other domains requiring pattern recognition and problem-solving capabilities.

1. **Introduction**

The profound challenge of solving incomplete Sudoku puzzles has traditionally been approached through various heuristic algorithms, each with its own set of limitations and complexities. This paper introduces a novel application of Convolutional Neural Networks (CNNs) to this domain, inspired by an assignment from the CNED institution. The assignment's core objective was to *develop software capable of efficiently solving Sudoku puzzles*, a task that has intrigued and challenged enthusiasts and researchers alike. The inception of this research can be attributed to Grégory Stehlík Buffard, who posited the innovative hypothesis that CNNs could be adept at solving incomplete Sudoku puzzles. This hypothesis marks a departure from conventional heuristic-based methods, aiming instead for a solution that leverages the pattern recognition prowess of CNNs. The primary goal of this endeavor is to realize a high-efficiency system that solves Sudoku puzzles accurately and swiftly, circumventing the constraints and intricacies inherent in heuristic algorithms.

*Fig.: A graph comparing the speed of resolution by CNN and heuristic algorithms*

Despite the widespread interest in Sudoku as a problem-solving exercise, the application of CNNs in this context remains largely unexplored. This research venture is, to the best of our knowledge, pioneering in its attempt to harness deep learning for Sudoku puzzle resolution. The exploration is deeply personal and driven by a quest to uncover the untapped potential of neural networks in logical puzzle solving, an area ripe for academic and practical innovation. This paper outlines the methodology adopted in training a CNN model on a dataset comprising millions of Sudoku puzzles, detailing the model architecture, training process, and validation strategy employed. Through rigorous experimentation and analysis, we aim to demonstrate the model's capability to fill incomplete puzzles with a high degree of accuracy, thereby offering a compelling alternative to traditional solving methods.

Our research not only contributes to the field of puzzle solving but also extends the applicability of CNNs to problems characterized by logical and spatial constraints. By documenting our findings and insights, we hope to inspire further investigation and development in this intriguing intersection of deep learning and recreational mathematics.

1. **Related Work**