

# Supervised Learning Programming Assignment - 1

## Datasets to be used

---

- Connect-4
- Solar Flares

## Background Description

---

### Connect-4

Connect-4 is a two-player strategy game, introduced by Milton Bradley (and now owned by Hasbro) in 1974. The objective, for each player, is to be the first player to have tokens (or discs) that are in a row, either vertically, horizontally or diagonally. The full list of rules can be downloaded at: [Hasbro Official Rules - \[https://instructions.hasbro.com/api/download/A5640\\_en-my\\_connect-4-game.pdf\]](https://instructions.hasbro.com/api/download/A5640_en-my_connect-4-game.pdf)

The game can end in either a `win`, `loss`, or a `tie` (sometimes also called stalemate or draw).

Check out the accompanying data-description (connect-4-data-desc.pdf) to understand what the dataset contains.

*Task:* Perform supervised learning to predict whether a given position ends in a `win`, `loss`, or a `tie` for **Player 1**. Do this task in a notebook called `connect4.ipynb`

---

### Solar Flares

Solar flares are large eruptions on the surface of the sun. Though the mechanisms are not fully understood, flares typically occur in regions with complex magnetic field line arrangements such as those around sunspots. Minor sun flares are relatively common and their high-energy radiation and particle streams have little to no effect on Earth. Extreme sun flares are more rare, but considerably more disruptive to communications. This dataset consist of descriptions of regions with sunspots recorded in 1978, together with the number of flares originating from these regions. Note that we here focus on predicting the number of C-class flares, as other classes are considerably more rare, and the number of instances would make it impossible to predict such flares.

Check out the accompanying data-description (in YAML file - solar-flares-metadata.yaml) to understand the attributes inside the dataset.

*Task:* Perform supervised learning to predict the **number of C-class flares**. Do this task in a notebook called `solarFlares.ipynb`

---

### Comparison Task

The `Connect-4` dataset requires classification, while the `Solar Flares` dataset requires regression.

In the notebook for *Solar Flares*, at the end, create a section called **Reflection**. Based on the performance of the techniques that you used, reflect on whether a particular technique is better than all others, regardless of dataset or kind of task. If so, why? If not, why not? Refer to evidence from your own notebooks to justify your answer.

---

## Submission and Grading

---

*Submission:* one Jupyter Notebook per dataset containing the learning task.

*Deadline:* 27-October-2025 (this can be superceded by announcements or new deadline on Brightspace by the module coordinator)

The accompanying rubric will be used for grading your Jupyter Notebook. Not all aspects of data cleaning may apply to you. It is however, *essential*, that you address all aspects in your Jupyter Notebook. You should discuss why (or why not) you perform each aspect of learning. The discussion is as important as the code, and lack of discussion will be graded *downwards* appropriately.

---

---