Datathon Problem Statement: Cheating Detection in Online Gaming

Overview

Cheating in online gaming has become a significant issue, negatively impacting fair play and the gaming experience. Modern games integrate various security and anti-cheat mechanisms to detect suspicious activities, yet cheaters constantly evolve their methods to bypass detection. In this datathon, participants are tasked with developing a machine learning model to identify players who may be engaging in unfair play using a dataset that contains system attributes, behavioral patterns, and security configurations.

Objective

The goal of this challenge is to build a robust classification model that can accurately predict whether a player is engaging in cheating activities. The dataset provided includes a wide range of features, including system information, security settings, player behavior, and in-game telemetry. Your task is to use this data to develop a predictive model that can distinguish between normal players and suspected cheaters.

Dataset Description

The dataset consists of anonymized data points representing individual gaming sessions. Each row corresponds to a player's gaming environment, security configurations, and system attributes. The dataset includes the following feature groups:

- ## General Identifiers
    - PlayerID - Unique identifier for an individual player.
    - GameVersion - Version of the game client installed on the player's system.
    - GameEngineVersion - Version of the game engine running on the player's system.
    - AntiCheatVersion - Version of the installed anti-cheat software.
    - SecuritySignatureVersion - Version of the security signature database used for cheat detection.
    - BetaTesterFlag - Indicates whether the player is a beta tester.
    - RealTimeProtectionStatus - Status of real-time cheat protection (if applicable).
    - PassiveModeFlag - Indicates whether the anti-cheat software is running in passive mode.
    - GameClientID - Unique identifier for the game client installed on the machine.
    - AntiCheatStatus - Identifier for the specific configuration of a player's anti-cheat settings.
    - SecurityToolsInstalled - Number of security/anti-cheat tools installed on the player's machine.
    - SecurityToolsEnabled - Number of security/anti-cheat tools actively running.
    - SecureHardwareFlag - Indicates whether the player's machine has a secure hardware module enabled.

- **Player & Location Attributes**
  - PlayerCountryID - Unique identifier for the country where the player is located.
  - PlayerCityID - Unique identifier for the city where the player is located.
  - ISP_ID - Unique identifier for the player's internet service provider.
  - GeoRegionID - Unique identifier for the geographic region the player is in.
  - LanguageSetting - English name of the player's system locale.
- **System Information**
  - GamingPlatform - The platform type (PC, console, etc.) the player is using.
  - CPUArchitecture - The processor architecture of the player's system.
  - OSVersion - The version of the operating system running on the player's system.
  - OSBuildVersion - The build version of the operating system.
  - OSFeatureSet - Product suite mask for the operating system version.
  - OSReleaseType - The sub-release version of the operating system.
  - OSBuildDetails - Detailed version information about the OS build.
  - OSEdition - The edition type of the player's operating system (e.g., Home, Pro).
- **Security & Protection**
  - AntiCheatProtectionEnabled - Indicates whether an active and up-to-date anti-cheat system is enabled.
  - AutomaticDataSharing - Indicates whether the player has opted into automatic data sharing.
  - CheatingDetectionMode - Mode in which cheating detection is enabled.
  - SecureModeEnabled - Indicates whether the system is running in a restricted security mode.
  - EmbeddedBrowserVersion - Version of the embedded browser within the game client.
  - FraudDetectionStatus - Status of in-game fraud detection mechanisms.
  - FirewallEnabled - Indicates whether the player's firewall is enabled.
  - UserAccessControlLevel - Reports the level of user access control enabled on the player's system.
- **Hardware Attributes**
  - DeviceType - Classification of the device based on hardware characteristics (e.g., laptop, desktop).
  - DeviceFamily - The general type of device (e.g., gaming PC, workstation).
  - DeviceManufacturerID - Unique identifier for the manufacturer of the player's device.
  - DeviceModelID - Unique identifier for the specific model of the player's device.
  - CPUCoreCount - The number of logical CPU cores in the player's system.
  - CPUManufacturerID - Unique identifier for the CPU manufacturer.
  - CPUModelID - Unique identifier for the specific CPU model.
  - CPUPerformanceClass - Categorization of CPU performance into high/medium/low tiers.
  - StorageCapacity - Total storage capacity of the primary drive in MB.
  - StorageType - Type of storage used (HDD or SSD).

- SystemPartitionSize - Size of the partition where the operating system is installed.
- HasDiskDrive - Indicates whether the player's device has an optical disk drive.
- RAMSize - The total physical RAM installed on the player's system.
- ChassisType - Type of computer chassis (e.g., tower, laptop, all-in-one).
- ScreenSize - Diagonal screen size of the primary display in inches.
- ScreenResolutionWidth - Horizontal resolution of the primary display in pixels.
- ScreenResolutionHeight - Vertical resolution of the primary display in pixels.
- PowerMode - Preferred power management profile based on the device type.
- Battery Attributes
- BatteryType - Type of internal battery used in the player's device.
- BatteryChargeCycles - The number of charge cycles the internal battery has undergone.

● Operating System & Update Info

- OSVersionDetails - Detailed version information of the operating system.
- OSArchitecture - The system architecture (e.g., 64-bit, ARM).
- OSDevelopmentBranch - Development branch of the OS (e.g., Insider builds).
- OSBuildNum - OS build number extracted from the system version.
- OSBuildRev - Revision number of the OS build.
- OSEditionName - User-friendly name of the OS edition.
- OSLicenseType - License type of the operating system (e.g., Retail, OEM).
- OSInstallationType - Description of how the OS was installed (e.g., clean install, upgrade).
- OSInstallLanguage - Language setting at the time of OS installation.
- OSUILanguage - Language used in the user interface of the operating system.
- OSAutoUpdateSettings - User-defined Windows Update settings.
- PortableOSFlag - Indicates whether the OS is booted via a portable medium (e.g., USB).

● License & Genuine Check

- GameLicenseStatus - Indicates whether the game license is valid and genuine.
- LicenseType - Specifies whether the game is using a retail or volume license.
- InternalTestingFlag - Indicates whether the device is being used for internal testing.
- BetaTestingDisabled - Indicates if the player has disabled participation in beta testing.
- BetaTestingRing - The testing ring the player is subscribed to (e.g., alpha, beta).
- EarlyAccessProgram - Indicates whether the player is opted into early access builds.

● Firmware & Boot Security

- FirmwareManufacturerID - Unique identifier for the firmware manufacturer.
- FirmwareVersionID - Unique identifier for the firmware version.
- SecureBootEnabled - Indicates whether Secure Boot is enabled in BIOS.
- WIMBootFlag - Indicates if the OS uses WIMBoot (Windows Image Boot).

    - VirtualMachineFlag - Identifies whether the system is running on a virtual machine.
    - TouchScreenFlag - Indicates whether the device has a touchscreen.
    - StylusSupportFlag - Indicates whether the device supports stylus input.
    - AlwaysOnFlag - Indicates whether the device supports Always-On, Always-Connected mode.
- **Player & Region Data**
    - GamerFlag - Indicates whether the device is categorized as a gaming device.
    - PlayerRegionID - Unique identifier for the region where the player is located.
    - CheatingFlag - Target variable indicating whether the player has been flagged for cheating.

Problem Statement

Your task is to develop a classification model that can predict whether a given player is likely to be cheating based on the provided attributes. The key challenges include:

- **High Class Imbalance**: The number of cheaters is expected to be significantly lower than legitimate players.
- **Noisy Features**: Some system attributes may not be directly relevant to cheat detection but could introduce noise.
- **Obfuscated Data**: The feature names have been anonymized to prevent direct inference about their origin.

Evaluation Metrics

Models will be evaluated based on **Accuracy.**

Submission Guidelines

Participants must submit the following:

1. **Predicted labels on a test dataset** (to be provided during the competition).
2. **Jupyter Notebook** (a single jupyter notebook to be submitted at the end for the best submission).

Submission Link :- https://www.kaggle.com/t/495f3a1c1b5a498d9c8a49351034cda7

Conclusion

This datathon presents an exciting opportunity to tackle a real-world problem in gaming security. Participants are encouraged to experiment with different approaches, leverage domain knowledge, and develop innovative solutions that enhance fair play in online gaming environments. Good luck!