

# 사이버보안 AI-빅데이터 챌린지 2022

## 대회안내서

주최기관



과학기술정보통신부

주관기관



한국인터넷진흥원

후원사



# CONTENTS

1 챌린지 플랫폼 접속 및 회원가입

2 A트랙 | AI기반 악성코드 위협 분류

3 B트랙 | AI기반 네트워크 공격 분류

4 C트랙 | 사이버보안 사회문제 해결 아이디어 공모

# CONTENTS

1 챌린지 플랫폼 접속 및 회원가입

2 A트랙 | AI기반 악성코드 위협분류

3 B트랙 | AI기반 네트워크 공격분류

4 C트랙 | 사이버보안 사회문제 해결 아이디어 공모



챌린지 플랫폼을 통해 데이터셋 배포  
각 팀별 팀장은 회원가입 후 대회 정보 확인 및 데이터셋 다운로드, 제출물 제출, 공지사항 확인 가능

## 경진대회 홈페이지

**사이버보안 AI·빅데이터 챌린지 2022**

운영트랙

트랙A	트랙B	트랙C
AI기반 악성코드 위협 분류	AI기반 네트워크 공격 분류	사이버보안 사회문제 해결 아이디어 공모
악성코드 특성을 분석하여 고성능의 악성코드 위협타입을 분류하는 AI모델을 개발하라!	네트워크 보안장비로 탐지된 공격로그 분석을 통해, 고성능의 네트워크 공격유형을 분류하는 AI모델을 개발하라!	최근까지 이슈가 된 보안위협에 따른 사회문제를 제시하고, AI·빅데이터를 활용해 이를 해결할 수 있는 아이디어를 제시하라!

## CZON(Challenge ON) 챌린지 플랫폼

**사이버보안 AI 빅데이터 챌린지 2022**

2022. 09. 02.(금) ~ 2022. 12. 01(목)

27,000,000 원

과학기술정보통신부 | 한국인터넷진흥원

참가신청 시작 D-24

참가시작전

**회원가입**

CZON 약관 동의

CZON을 이용하기 위해 약관에 동의하십시오.

- ☑ 모든 약관에 동의합니다.
- ☑ 이용약관에 동의 (필수) ☐
- ☑ 개인정보 취급방침에 동의 (필수) ☐
- ☑ 마케팅 수신정보 제공에 동의 (선택) ☐

다음으로

회원가입

**데이터셋 다운로드**

대회소개 데이터 자료실 팀 제출 리더보드 공지사항

공동 데이터 예선 데이터

데이터 다운로드

데이터셋 다운로드

# CONTENTS

1 챌린지 플랫폼 접속 및 회원가입

2 A트랙 | AI기반 악성코드 위협 분류

3 B트랙 | AI기반 네트워크 공격 분류

4 C트랙 | 사이버보안 사회문제 해결 아이디어 공모

구분	날짜	대회방식	시간		내용
예선	'22.10.20(목)	예선셋 배포	10:00 ~ 24:00	840'	
	'22.10.21(금)	온라인 기술경연	9:30 ~ 10:00	30'	예선셋 압축 패스워드 공개
			10:00 ~ 17:00	420'	경연진행
본선	'22.11.16(수)	본선셋 배포	10:00 ~ 24:00	840'	
	'22.11.17(목)	온라인 기술경연	9:30 ~ 10:00	30'	본선셋 압축 패스워드 공개
			10:00 ~ 17:00	420'	경연진행
	'22.11.18(금)	오프라인 발표평가	13:00 ~ 17:00	240'	7개팀 발표평가

챌린지 플랫폼을 통해 데이터셋 배포

## 데이터셋 배포 일정

- 학습셋 오픈 : '22. 9. 29(목) 10:00 ~ '22. 10. 14(금) 18:00
- 예선셋 오픈 : '22. 10. 20(목) 10:00 ~ 24:00
- 본선셋 오픈 : '22. 11. 16(수) 10:00 ~ 24:00

## 데이터셋 구성

- 학습셋) 악성코드 원본 30,000건 + 라벨파일(csv)
- 예선셋) 악성코드 원본 20,000건
- 본선셋) 악성코드 원본 20,000건

※ 예·본선 데이터셋은 데이터셋 구축 및 검증 과정에서 일부 변경될 수 있음

The screenshot shows the ZON Challenge 2022 website. At the top, there's a navigation bar with 'ZON CHALLENGE 2022', '챌린지', '공지사항', and '교육'. On the right, there's a user profile section with 'SYSNAR' and '로그아웃'. The main banner features the title '사이버보안 AI 빅데이터 챌린지 2022' and a '참가 신청 시작 D-24' badge. Below the banner, there's a list of details: '사이버보안 AI 응용기술의 혁신을 선도할 아이디어와 AI 우수 인재발굴', '2022. 09. 02.(금) ~ 2022. 12. 01(목)', '27,000,000 원', and '과학기술정보통신부 | 한국인터넷진흥원'. A '참가시작전' button is visible on the right. Below the banner, there's a navigation menu with '대회소개', '데이터', '자료실', '팀', '제출', '리더보드', and '공지사항'. Under the '데이터' tab, there are two sub-tabs: '공통 데이터' and '예선 데이터'. At the bottom, there's a large red button labeled '데이터 다운로드'.

## 양식에 맞게 결과물 작성 후 제출

### 결과물 작성 ▲분류결과 ▲기술보고서 ▲발표자료

- (분 류 결 과) ID: 파일명 | class: 위협타입+패밀리 | 양식: CSV | 파일명: 팀명\_track\_a\_answer | 제출 : 예·본선
- (기술보고서) 배포한 양식에 포함된 필수항목에 대한 내용 작성 | 파일명: 팀명\_track\_a\_report | 제출: 예선
- (발 표 자 료) 기술보고서 내용을 요약해 발표 15분 분량으로 작성 | 파일명: 팀명\_track\_a\_final | 제출: 본선

※ 오프라인 심사장에 팀원 참석 여부는 코로나19 상황을 고려해 추후 결정 후 공지 예정

<분류결과 제출 샘플>

	ID	Class	
1	ID	위협타입	패밀리
2	파일명	A	a
3			
4			
5			



## 챌린지 플랫폼의 제출 탭을 통해 결과물 제출

대회소개

데이터

자료실

팀

제출

라더보드

공지사항

제출횟수 현황 3/5

클릭하여 제출할 파일을 선택해주세요.

선택된 파일

파일명	확장자	용량

제출하기

제출정보  
private 점수는 대회 종료 후 공개됩니다.

제출순서	제출자	제출일자	파일명	public	private	기타 메시지
24	SYSNAR	2022-08-23T17:14:29.607712	example_team3csv	-	-	채점대기
23	SYSNAR	2022-08-23T17:02:06.647148	example_team3csv	-	-	채점대기
22	SYSNAR	2022-08-22T14:48:18.576135	test_answer (1).csv	1	-	채점완료
21	SYSNAR	2022-08-19T10:00:34.478165	test_answer (1).csv	1	-	채점완료
20	SYSNAR	2022-08-16T16:57:17.348619	test (1).csv	-	-	채점오류

결과물 제출 탭

제출할 파일을 찾아서 선택 혹은  
드래그&드롭하여 업로드

선택한 파일 리스트 확인

제출하기 버튼을 클릭하여 제출

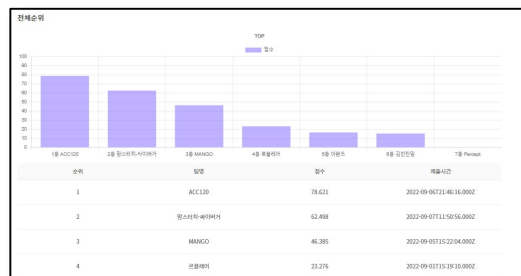
※ 한 번 제출한 제출물은 수정이 불가하오니, 유의하시어 업로드 바랍니다.

횟수별 제출물 정보 확인

(예선)은 온라인 기술경연, (본선)은 온라인 기술경연+오프라인 발표평가로 진행되며, 리더보드 및 공지사항은 경연 플랫폼 CZON(챌린지온)을 이용

## 예·본선 기술경연 진행방식

- (제출기회) 경연시간 內 [분류결과]는 총 5회까지 제출 가능하며, 최종 분류정확도는 마지막 제출한 결과로 반영  
- [기술보고서]는 예선 마감시간 전까지 경연페이지에 제출 업로드  
※ 제출 마감시간 이후에는 제출 받지 않으니, 시간 엄수 必
- (분류정확도 확인) 리더보드를 통해 팀별 실시간 분류정확도 비교경쟁
- (본선진출팀 발표) '22. 11. 1(화)중 누리집을 통해 발표예정  
※ 예선 종료 후, 평가위원회를 통한 서면평가 및 치팅검증 예정(약 5일)



CZON 리더보드(예시)

## 본선 발표경연 진행방식

- (결과물 제출) '22.11.17(목), 17:00까지 경연페이지에 발표자료 제출 업로드
- (진행방식) 발표자만 오프라인 심사장에 참석하여 발표 평가  
- 발표는 최종 발표자료 제출의 역순으로, 팀별 발표시간 별도 공지예정  
※ 오프라인 심사장에 팀원 참석 여부는 코로나19 상황을 고려해 추후 결정 후 공지 예정
- (평가시간) 15분 발표 + 10분 질의 (±5분)



CZON 공지사항

평가지표는 평가위원회 의견에 따라 공개 가능한 수준으로만 공지드립니다.

## 평가지표

- (예선) 분류정확도(60%) + 순위별 차등점수\* (10%) + 서면평가(30%)
- (본선) 분류정확도(80%) + 발표평가(20%)

\* 순위별 차등점수: 1~3등 : 10점 | 4~6등 : 8점 | 7~9등 이하 6 | 10등 이하 4점

※ 기술보고서는 서면평가 및 치팅검증에 활용

- (분류정확도) Precision with Macro-Average (별첨)

## 평가방법

- 대회기간 제공되는 대회용 데이터셋은 Virus Total 등 검색엔진에 활용 제한  
※ 검색엔진 활용 등의 부정행위 발견 시, 수상에서 제외될 수 있습니다.
- (블라인드 평가) 평가위원회는 팀정보를 모르는 상태에서 평가진행
- 평가위원(8인 내외) 평가점수 중 최고 및 최저 점수 제외 후 최종 점수 도출

## 평가항목

※ 평가위원회 의견에 따라 각 평가항목에 대한 세부 평가기준 및 배점 비공개

NO.	평가항목
1	데이터 구성의 이해
2	feature 엔지니어링
3	데이터 전처리기술
4	AI 기술수준
5	기타 배점(독창성 등)

# CONTENTS

1

챌린지 플랫폼 접속 및 회원가입

2

A트랙 | AI기반 악성코드 위협분류

3

B트랙 | AI기반 네트워크 공격 분류

4

C트랙 | 사이버보안 사회문제 해결 아이디어 공모

구분	날짜	대회방식	시간		내용
예선	'22.10.20(목)	예선셋 배포	10:00 ~ 24:00	840'	
	'22.10.21(금)	온라인 기술경연	9:30 ~ 10:00	30'	예선셋 압축 패스워드 공개
			10:00 ~ 17:00	420'	경연진행
본선	'22.11.16(수)	본선셋 배포	10:00 ~ 24:00	840'	
	'22.11.17(목)	온라인 기술경연	9:30 ~ 10:00	30'	본선셋 압축 패스워드 공개
			10:00 ~ 17:00	420'	경연진행
	'22.11.18(금)	오프라인 발표평가	13:00 ~ 17:00	240'	7개팀 발표평가



챌린지 플랫폼을 통해 데이터셋 배포

## 데이터셋 배포 일정

- 학습셋 오픈 : '22. 9. 29(목) 10:00 ~ '22. 10. 14(금) 18:00
- 예선셋 오픈 : '22. 10. 20(목) 10:00 ~ 24:00
- 본선셋 오픈 : '22. 11. 16(수) 10:00 ~ 24:00

## 데이터셋 구성

- 학습셋) 공격 로그데이터 + 라벨파일(csv) 50,000건
- 예선셋) 공격 로그데이터 10,000건
- 본선셋) 공격 로그데이터 10,000건

※ 예·본선 데이터셋은 데이터셋 구축 및 검증 과정에서 일부 변경될 수 있음

The screenshot shows the ZON Challenge 2022 website. At the top, there's a navigation bar with 'ZON CHALLENGE 2022', '챌린지', '공지사항', and '교육'. On the right, there's a user profile section with 'SYSNAR' and a '로그아웃' button. The main banner features the title '사이버보안 AI 빅데이터 챌린지 2022' and a '참가 신청 시작 D-24' badge. Below the title, it states the purpose: '사이버보안 AI 응용기술의 혁신을 선도할 아이디어와 AI 우수 인재발굴'. The dates are '2022. 09. 02.(금) ~ 2022. 12. 01(목)' and the prize is '27,000,000 원'. It also mentions '과학기술정보통신부 | 한국인터넷진흥원'. A '참가시작전' button is visible. Below the banner, there's a navigation menu with '대회소개', '데이터', '자료실', '팀', '제출', '리더보드', and '공지사항'. Under '데이터', there are two tabs: '공통 데이터' and '예선 데이터'. A large '데이터 다운로드' button is at the bottom.

챌린지 플랫폼의 제출 탭을 통해 결과물 제출

## 결과물 작성 ▲분류결과 ▲기술보고서 ▲발표자료

- (분 류 결 과) ID: 파일명 | class: 공격유형 | 양식: CSV | 파일명: 팀명\_track\_b\_answer | 제출: 예·본선
- (기술보고서) 배포한 양식에 포함된 필수항목에 대한 내용 작성 | 파일명: 팀명\_track\_b\_report | 제출: 예선
- (발 표 자 료) 기술보고서 내용을 요약해 발표 15분 분량으로 작성 | 파일명: 팀명\_track\_b\_final | 제출: 본선

※ 오프라인 심사장에 팀원 참석 여부는 코로나19 상황을 고려해 추후 결정 후 공지 예정

<분류결과 제출 샘플>

	A	B
1	id	class
2	A	recon
3		
4		

## 챌린지 플랫폼의 제출 탭을 통해 결과물 제출

대회소개

데이터

자료실

팀

제출

라더보드

공지사항

제출횟수 현황 3/5

클릭하여 제출할 파일을 선택해주세요.

선택된 파일

파일명	확장자	용량

제출하기

제출정보  
private 점수는 대회 종료 후 공개됩니다.

제출순서	제출자	제출일자	파일명	public	private	기타 메시지
24	SYSNAR	2022-08-23T17:14:29.607T12	example_team3csv	-	-	채점대기
23	SYSNAR	2022-08-23T17:02:06.647T148	example_team3csv	-	-	채점대기
22	SYSNAR	2022-08-22T14:48:18.576T135	test_answer (1)csv	1	-	채점완료
21	SYSNAR	2022-08-19T10:00:34.478T165	test_answer (1)csv	1	-	채점완료
20	SYSNAR	2022-08-16T16:57:17.348T619	test (1)csv	-	-	채점오류

결과물 제출 탭

제출할 파일을 찾아서 선택 혹은  
드래그&드롭하여 업로드

선택한 파일 리스트 확인

제출하기 버튼을 클릭하여 제출

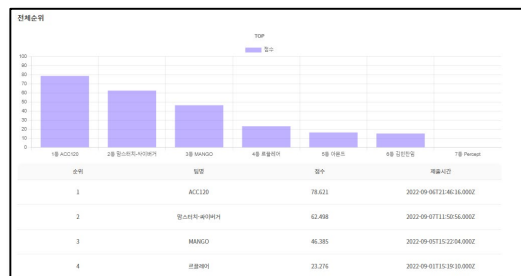
※ 한 번 제출한 제출물은 수정이 불가하오니, 유의하시어 업로드 바랍니다.

횟수별 제출물 정보 확인

(예선)은 온라인 기술경연, (본선)은 온라인 기술경연+오프라인 발표평가로 진행되며, 리더보드 및 공지사항은 경연 플랫폼 CZON(챌린지온)을 이용

## 예·본선 기술경연 진행방식

- (제출기회) 경연시간 內 [분류결과]는 총 5회까지 제출 가능하며, 최종 분류정확도는 마지막 제출한 결과로 반영  
- [기술보고서]는 예선 마감시간 전까지 경연페이지에 제출 업로드  
※ 제출 마감시간 이후에는 제출 받지 않으니, 시간 엄수 必
- (분류정확도 확인) 리더보드를 통해 팀별 실시간 분류정확도 비교경쟁
- (본선진출팀 발표) '22. 11. 1(화)중 누리집을 통해 발표예정  
※ 예선 종료 후, 평가위원회를 통한 서면평가 및 치팅검증 예정(약 5일)



CZON 리더보드(예시)

## 본선 발표경연 진행방식

- (결과물 제출) '22.11.17(목), 17:00까지 경연페이지에 발표자료 제출 업로드
- (진행방식) 발표자만 오프라인 심사장에 참석하여 발표 평가  
- 발표는 최종 발표자료 제출의 역순으로, 팀별 발표시간 별도 공지예정  
※ 오프라인 심사장에 팀원 참석 여부는 코로나19 상황을 고려해 추후 결정 후 공지 예정
- (평가시간) 15분 발표 + 10분 질의 (±5분)



CZON 공지사항

평가지표는 평가위원회 의견에 따라 공개 가능한 수준으로만 공개 가능

## 평가지표

- 예선) 분류정확도(90%) + 서면평가(10%)
- 본선) 분류정확도(70%) + 발표평가(30%)  
※ 기술보고서는 서면평가 및 치팅검증에 활용
- (분류정확도) Precision with Macro-Average (별첨)

## 평가방법

- (블라인드 평가) 평가위원회는 팀정보를 모르는 상태에서 평가진행
- 평가위원(8인 내외) 평가점수 중 최고 및 최저 점수 제외 후 최종 점수 도출

## 평가항목

※ 평가위원회 의견에 따라 각 평가항목에 대한 세부 평가기준 및 배점 비공개

NO.	평가항목
1	데이터 구성의 이해
2	feature 엔지니어링
3	데이터 전처리 기술
4	AI 기술수준
5	기타 배점(독창성 등)



# CONTENTS

1 챌린지 플랫폼 접속 및 회원가입

2 A트랙 | AI기반 악성코드 위협 분류

3 B트랙 | AI기반 네트워크 공격 분류

4 C트랙 | 사이버보안 사회문제 해결 아이디어 공모

구분	날짜	대회방식	시간		내용
예선	~ '22.10.21(금)	공모서 제출	~ '22.10.21(금) 17:00		서면평가 10.24(월) ~ 10.28(금)
본선	'22.11.18(금)	오프라인 발표평가	13:00 ~ 17:00	240'	7개팀 발표평가

## 챌린지 플랫폼을 통해 결과물 제출

### 예선 진행방식

- (제출기회) 제출기한 내 마지막에 접수 된 공모문서로 서면평가 실시
- (서면평가) 예선 종료 후, 평가위원회를 통한 공모서 서면평가 진행  
※ 서면평가 진행 중, 부정행위(아이디어 도용 등) 발견 시 수상에서 제외
- (본선진출팀 발표) '22. 11. 1(화)중 누리집을 통해 발표예정

대회소개	데이터	자료실	팀	제출	리더보드	공지사항
공지사항						
2	[공지] CZON 이용약관 수정 안내 2022-09-25(15:58:39)관리자					
1	이벤트 모든 이벤트 2022-09-20(15:58:39)관리자					

CZON 공지사항

### 본선 발표경연 진행방식

- (결과물 제출) '22.11.17(목), 17:00'까지 경연페이지에 발표자료 제출 업로드
- (진행방식) 발표자만 오프라인 심사장에 참석하여 발표 평가  
- 발표는 최종 발표자료 제출의 역순으로, 팀별 발표시간 별도 공지예정  
※ 오프라인 심사장에 팀원 참석 여부는 코로나19 상황을 고려해 추후 결정 후 공지 예정
- (평가시간) 15분 발표 + 10분 질의 (±5분)

평가지표는 평가위원회 의견에 따라 공개 가능한 수준으로만 공개 가능

## 평가지표

- 예선) 서면평가 (100%)
- 본선) 예선결과(80%) + 발표평가(20%)

※ 본선에서는 예선의 서면평가 결과를 80%로 반영

## 평가방법

- (블라인드 평가) 평가위원회는 팀정보를 모르는 상태에서 평가진행
- 평가위원(8인 내외) 평가점수 중 최고 및 최저 점수 제외 후 최종 점수 도출

## 평가항목

※ 평가위원회 의견에 따라 각 평가항목에 대한 세부 평가기준 및 배점 비공개

NO.	평가항목
1	(시의성) 아이디어 제안배경 및 필요성
2	(실현성) 제안내용의 구현 및 실현 가능성
3	(파급성) 기대효과 및 사회적 파급력
4	(창의성) 아이디어의 독창성·창의성

The background is a dark blue field with a glowing blue grid pattern that recedes into the distance. Two faint, light blue shield icons are positioned on the left and right sides of the central text.

# 사이버보안 AI-빅데이터 챌린지 2022

감사합니다





# [별첨] 분류정확도 산정방식

## Confusion Matrix

		Prediction	
		Positive	Negative
Actual	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

## Precision

- Positive로 예측한 비율 중 진짜 Positive인 비율
- $Precision_i$ 는 클래스 i에 대한 Precision 값

$$Precision_i = \frac{True\ Positive_i}{True\ Positive_i + False\ Positive_i}$$

## Precision with Macro-Average

- k개의 클래스에 대해, 각 클래스 별 Precision 합이 평균

$$Average = \sum_{i=1}^k \frac{1}{k} Precision_i$$