

연세대학교 통계 데이터 사이언스 학회 ESC

23-2 SUMMER WEEK 1

1. 예수 스테디



예습 스터디

예습 스터디에서 노선에 기록한 자료는 추후 ESC vlog 계정에 공개 업로드 될 예정입니다! 따라서 예습 스터디에서 이용하신
교재 및 외부 자료의 출처 표기에 신경 써주시면 감사하겠습니다.



1.2.3 Data Error and Computational Error

x : true value of the input, $f(x)$: desired true result

\hat{x} : inexact input, $\hat{f}(x)$: approximation to the function

$$\begin{aligned}\text{Total error} &= \hat{f}(\hat{x}) - f(x) \\ &= (\hat{f}(\hat{x}) - f(\hat{x})) + (f(\hat{x}) - f(x)) \\ &= \text{computational error} + \text{propagated data error}.\end{aligned}$$

- 두 번째 줄 식
 - 같은 input에 대한 exact & approximate function의 차 = computational error
 - exact function 값에 대한 차이 (\hat{x} & x) = propagated data error



Most realistic problems are multidimensional, but for simplicity we will consider only one-dimensional problems in this chapter; extension of the definitions and results to higher dimensions is straightforward, usually requiring only replacement of absolute values by appropriate *norms* (see Section 2.3.1). A typical problem in one dimension can be viewed as the computation of the value of a function, say $f: \mathbb{R} \rightarrow \mathbb{R}$, that maps a given input value to an output result. Denote the true value of the input by x , so that the desired true result is $f(x)$. Suppose that we must work with inexact input, say \hat{x} , and we can compute only an approximation to the function, say \hat{f} . Then, using the standard mathematical trick of adding and subtracting the same quantity so that the total is unchanged, we have

$$\begin{aligned}\text{Total error} &= \hat{f}(\hat{x}) - f(x) \\ &= (\hat{f}(\hat{x}) - f(\hat{x})) + (f(\hat{x}) - f(x)) \\ &= \text{computational error} + \text{propagated data error}.\end{aligned}$$

교재에 있는 그림이나 복잡한 수식의 경우 캡처해도 괜찮지만(LaTeX로 쓰면 더 좋습니다), 간혹 책 글까지 복사해서 붙여넣기 하는 경우가 있습니다. 대부분의 학회원들이 정성 들여 예습 스터디에 참여해주시는 만큼 이러한 행위는 자제해주세요! 더 좋은 퀄리티의 ESC vlog를 만들어가는 과정에 모두가 동참할 수 있었으면 좋겠습니다 :)

2. 친해지길 바라



ESC 23-2 친바

날짜: 7월 22일 토요일

시간: 조별로 조율하기! (뒤풀이 & 식사도 조별로 진행)

활동: 볼링 & 보드게임 & 축구 보러 가기 중 택1 (축구: 서울 vs 인천 in 월드컵 경기장)

7월 13일, 14일 (이번주 목요일, 금요일) 이틀 간 활동 선호 조사 진행 (중복 투표 가능)

- 투표 결과를 기반으로 이후 조 편성 예정
- 친바에 참여 가능하신 분들은 하고 싶은 활동에 투표해주세요!
- 조 편성을 위해, 일부 시간만 참여가능한 경우 활동 투표와 함께 가능한 시간을 댓글로 달아주세요!

3. 출석 규칙



출석 규칙 개정

8월 연합 연사특강이 기존 8월 5일에서 5일, 12일, 26일 중 하루로 변경될 수 있음 (고려대 KUBIG 주최)
일정 변동의 가능성을 고려하여 출석 규칙을 일부 수정하고자 논의 중

(이전) 방학 세션 2회, 학기 세션 2회 이상 결석 시 제명 (=방학, 학기 각각 1결석 1지각 허용)

(수정) 방학 & 학기 세션 합쳐서 3회 결석까지 허용

이외에도 만약 (1) 8월 연합 연사특강의 일정이 변경되고 (2) 제명 위기의 학회원이 해당 특강에 결석하였을
경우 회장단 내 추가 논의를 통해 해당 날짜에 대한 출석 규칙을 완화할 예정

*친바, MT는 출석에 포함되지 않습니다!

4. 공모전 안내



Dataon 연구 데이터 분석 활용 경진대회

2023 DATA•AI 분석 경진대회

DATA•AI 기반의 과학•사회적 문제 해결 및 DNA 저변 확대

과학기술 문제 참가신청 바로가기

- 01.국가연구데이터통합플랫폼 DataON의 연구데이터에서 과학기술표준분류 예측
- 02.과학기술 논문 원문 정보로부터 주제 분류하기(Top3)
- 03.사상체질 판별 모델 개발
- 04.ScienceON 사용자 행동 예측
- 05.AWS 관측자료를 활용한 하천 유량 예측 모델 개발

사회현안 문제 참가신청 바로가기

- 01.A를 활용한 쓰레기 무단투기 예방
- 02.낙엽송, 상수리나무, 편백 조림목 생육예측모델 개발
- 03.BGP data를 활용한 이상 징후 탐지(BGP anomaly detection)
- 04.AI 영상 기술을 활용하여 PCB 표면 불량 판단
- 05.사과 품질 예측 모델 개발
- 06.기차 대차의 이상 소음원 위치 판별
- 07.지상 관측과 레이더 관측 자료를 활용한 주요 지점별 강우량 예측 모델 개발
- 08.뉴스와 소셜데이터 기반 이슈 분석 및 시각화
- 09.국회회의록 요약 모델 개발
- 10.한국어 과학기술 논문의 초록 문장분류
- 11.스마트팜 데이터를 이용한 작물 수확량 및 농가 수익 예측

부문	구분	수상자(팀)	상금 내역
문제발굴	-	과학기술 : 5명 사회현안 : 5명	· KISTI 원장상 · 각 상금 100만원
모델개발	대상	통합 1팀	· 국회의장상 · 상금 500만원
	최우수상	과학기술 : 1팀 사회현안 : 1팀	· 과기부 장관상 · 각 상금 300만원
	우수상	과학기술 : 2팀 사회현안 : 2팀	· 연구회 이사장상, 대전광역시장상 · 각 상금 200만원
	장려상	과학기술 : 3팀 사회현안 : 3팀	· KISTI 원장상, 국회도서관장상 · 각 상금 100만원

대회일정

- 대회공고 및 문제 발굴 : '23.5.30.(화) ~ 6.30.(금)
- 문제 공개 : '23.7.7(금)
- 워크숍(경진대회 설명회) : '23.7.13.(목)
- 참가신청 : '23.7.10.(월) ~ 7.28.(금)
 - 제출방법 : 경진대회 홈페이지 신청란 및 행사 사무국(별도 공지)
 - 관련 문의 : 행사 사무국(연락처 별도 공지)
- 1차 서류평가 및 결과발표 : '23.8.4.(금)
- 사례 및 모델 개발 기간 : '23.8.7.(월) ~ 10.20.(금)
- 발표자료 제출 : '23.10.20.(금)
- 최종발표 평가 : '23.10.26.(목) ~ 27.(금)
- 시상식 : '23.11.30.(목) (2023 미래연구정보포럼 행사)

참여 희망 시 내일 중으로 공지방에 올릴 수요 조사 카톡에 체크 표시, 직접 팀 구성(회장에게 사전 연락) or 회장단이 임의로 팀 구성하여 진행

5. 세션 준비 가이드



타임라인

파트 분배 - 노선 정리 - **예습 스터디 진행** - 발표 구성 - 발표 자료 제작
- **자체 리허설** - 학술부장 피드백 - 피드백 반영 - **발표**

발제 준비는 각 조의 역량에 따라 자율적으로 진행하면 되지만, 팀 단위 발제는 다들 생소하므로
30기 임원진이 직접 발제 준비를 해보며 참고할 만한 가이드라인을 제작하였음!

파트 분배

(이전) 예습 스터디와 본인의 발제 파트가 전혀 상관 없이 이루어짐

(현재) 팀 단위로 예습 스터디와 발제 준비 함께 진행하므로 동일한 파트를 처음부터 끝까지 맡을 수 있음

임원진은 범위를 3등분 후
각 파트를 2명씩 분담!

노선 정리 & 예습 스터디 진행

노선 정리를 꼼꼼히 해두면 발표 자료 제작에 드는 시간을 아낄 수 있음! 학술부장님께 조언 구하면 좋을 듯

뒷부분 담당자도 앞부분을 숙지해야 준비가 가능하므로, 노선 정리를 시간차를 두고 진행하는 것을 권장

(예시) 파트 2등분 or 4등분 후 앞의 두 명은 금요일, 뒤의 두 명은 토요일 내로 노선 정리 완성, 함께 읽어본 후 일요일에 예습 스터디 진행

발표 구성 & 발표 자료 제작

내용을 잘 숙지하는 것도 중요하지만, 좋은 발표를 위해 필요한 것이 몇 가지 더 있다.

- ① 어떻게 하면 학회원에게 내용을 쉽게 잘 전달할 수 있을지 고민하기
- ② 전체적인 발표 흐름이 유기적으로 연결되는지 확인하기
- ③ 노트이션, 표현 등을 일관성 있게 통일하기

예습 스터디 직후 논의를 통해 이를 수정 및 보완하는 작업을 반드시! 거쳐야 함

발표 자료의 경우 각자 예습 스터디에서 말았던 파트에 대한 PPT(글씨 & 자료만 첨부) 제작 후 이를 한 명의 팀원이 수합하는 방법을 통해 제작될 수 있음. 노선을 캡처해서 첨부하거나 PPT에 내용을 입력하는 두 방식이 있는데, 각각의 장단점이 있으므로 노선을 보기 좋게 신경 써서 정리하거나 PPT를 수합하기 좋게 만들어 넘겨줄 것!

전자는 PPT에 올릴 때 매우 편하지만 개별 내용 수정 or 강조가 어렵고, 후자는 성의가 좋고 편집이 쉬우나 PPT에 올리기 힘들다!
잘 정리해서 넘겨주지 않으면 자료 수합에 생각보다 매우 많은 시간이 든다 ㅠ ㅠ

자체 리허설 & 피드백 & 발표

학술부 세션은 따로 이루어지지 않지만, 팀의 발제자는 **자체 리허설**을 통해 준비한 **내용을 시간 내에 다 전달할 수 있을 지 테스트**해봐야 한다. 이를 토대로 분량 조절 및 내용 검토를 위해 학술부장과 피드백을 주고받는 시간을 갖게 된다! 이후 피드백을 반영하여 발표 자료를 완성하고 오는 목요일 세션에서 발표하면 끝~

공지, 쉬는 시간, 학술부장 & 감사님
피드백, 발표자의 긴장 등을
고려하여 발표가 시간 30분
이상으로 길어지는 것은 지양할 것!

공통 템플릿 안내

디자인 및 정리에 큰 시간을 쓰지 않도록 & 매 주차 발표가 통일성 있게 이어지도록 **공통 템플릿**을 제작하였음
내용 외의 부분은 수정 없이 사용하면 되고, 폰트는 **나눔스퀘어_ac**(기본, Bold, ExtraBold)만 사용! 설치 요망

역할 분담은 자유롭게 하기, 발표 당일
발표자는 조금 일찍 오는 것을 권장~

완성된 발표 자료는 PDF로
변환하여 세션 당일 오후 6시까지
조장이 공지방에 업로드하기!

연세대학교 통계 데이터 사이언스 학회 ESC 23-2 SUMMER WEEK(숫자)

제목 (영문)

[ESC 방학세션 (숫자)조] (이름) (이름) (이름) (이름)



(숫자).(소제목)

소제목 (영문)

페이지 주제 (영문)

내용 (정렬 자유롭게 전환)

