

SKills for AI & Data Analytics

SKADA 인증 시험

[상세 소개]

카드를 우측으로 넘겨
SKADA에 대해 알아보세요!



인증 Test는 어떻게 진행되나요?



▣ 문항 구성

객관식 (30점)

+

실습형 (70점)

*상세 구성 및 문제 예시는 뒷 페이지를 참고해주세요

⌚ 응시 소요 시간

4시간

☑ 합격 기준

70점 이상 합격

(Practitioner / Advanced Practitioner 동일)

▣ TEST 환경

mySUNI INFRA 시스템에서 응시

오프라인 시험장에서 클라우드 환경으로 이론/실습 테스트 진행

Practitioner는 어느 수준인가요?



전공/직무와 무관하게 Data 분석 관련 학습 및 1년 이상의 업무 적용 경험을 통해 정형 Data 분석 과제 (Data 추출-전처리/EDA-모델링)을 자체적으로 수행할 수 있는 Skill을 보유한 수준의 구성원입니다.

합격자 Profile 예시



김나라 님

① Profile

소속 회사 SK 하이닉스

소속 부서 NAND Etch 기술 관련 부서

담당 업무 공정 장비 유지 관리, 센서 Data 분석을 통해 변수에 대한 분석 및 유지/보수 판단

② Background

이수 과정

- ML Engineer ('23년)
- CDS ('21년)

Domain 경험 총 6년

- Etch 공정 장비 관리 6년 ('17년~현재)

Data Skills 총 6개월

- anomaly detection을 통한
센서 장비 고장 전 변곡점 파악 프로젝트

③ Data 분석 경험/프로젝트

주 사용 Data 유형 시계열 Data (공정 센서 Data)

Data 수집/추출 DB에 수집된 Data 직접 추출

Data Pre-processing Domain 지식을 기반으로
Data 분류를 단독 수행

Modeling 회귀, 이상 탐지 등 전반적인
분석 알고리즘 모델링을 단독 수행

Advanced Practitioner는 어느 수준인가요?



전공자 수준의 Data 분석 심화 학습 및 3년 이상의 업무 적용 경험을 통해
최신 ML/DL을 적용하여 복잡 다양한 데이터 분석 과제 (이미지, 시계열, 자연어)를
수행할 수 있는 Skill을 보유한 수준의 구성원입니다.

합격자 Profile 예시



이진석 님

① Profile

소속 회사 SK 텔레콤

소속 부서 Infra 관련 업무 부서

담당 업무 AI&DATA를 활용한 네트워크 품질개선,
이상징후 감지, 유지보수 업무 수행

② Background

이수 과정

- ML Engineer ('22년)
- Big Data 대학원 졸업

Domain 경험 총 14년

- 통신 네트워크 관리 10년
- ICT Infra Data 분석 4년

Data Skills 총 4년

- 통신망 품질 개선 프로젝트 등 AI 접목
- 통신 Infra 개선 과제 발굴/진행

③ Data 분석 경험/프로젝트

주 사용 Data 유형 시계열 Tabular

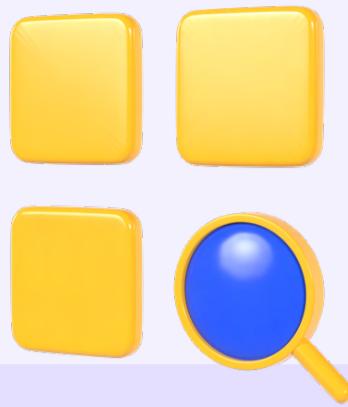
Data 수집/추출 Data lake에서 SQL 등 직접 추출

Data Pre-processing Feature Engineering 및
Data Frame 구축

Modeling

- 전반적인 ML 알고리즘 능숙 활용
- DL 활용 시계열 Data 분석 및
파라미터 튜닝 단독 수행

객관식 문제 유형은 어떻게 되나요?



30점 배점 객관식은 개념 확인, 수리연산, 코드 검증 등 선택형 문제로 구성됩니다.



개념확인형



ML 개념에 대해 맞는 설명을 찾거나
적절한 키워드를 선택하는 문제

수리연산형



ML의 수학적 의미를 이해하고
연산을 수행하는 문제

코드검증형



오류가 있는 코드를 디버깅하거나,
아키텍처/파라미터의 의미를 묻는 문제

시나리오형



주어진 시나리오에서
적절한 가설 또는 해석을 묻는 문제

[개념확인형] 문제 예시

Q. 문제 3

머신러닝에서 편향(bias)과 분산(variance)에 대한
설명으로 옳지 않은 것은?

- 편향(bias)과 분산(variance)은 반비례(trade-off)관계에 있다.
- 복잡도가 높은 모델을 사용할 경우, 분산
(variance)이 증가하여 과적합(overfitting)이 발생
가능성이 있다.

[수리연산형] 문제 예시

Q. 문제 7

다음 함수의 최소값을 구하기 위해 경사하강법
(Gradient Descent)를 사용하고자 한다. Initial
point ($x=4$)에서 learning rate가 0.01인 경사하강
법을 한 번 실행했을 때, 업데이트 된 x 값은?
 $f(x) = 2x^2 - 5$

- 3.71
- 3.73

주관식 문제 유형은 어떻게 되나요?



70점 배점

주관식은 모델링 실무와 유사하며, 2개 내외의 Data-set으로 구성됩니다.

Data-set

A

B



Section 1 [전처리]

Data Ingestion, EDA

Data 입수 및 특성 확인,
데이터 마이닝을 통한 결과 도출

Data Visualization

데이터 시각화 및 인사이트 도출

Data Preprocessing

데이터 전처리 관련 구현

Section 2 [모델링]

Baseline Modeling

모델 훈련, 평가 등 구현

Improving Performance

모델 성능 개선 관련 구현

Section 3 [고도화]

In-depth Modeling

특정 문제 상황 해결 혹은
제공된 수식 구현

Practitioner

평가 기준 및 시험 범위는 어떻게 되나요?



주요 역량	평가 기준	범위
AI Math	중심경향 기초통계량과 산포도 등을 이해하고 계산할 수 있는가	Mean, Median, Mode Standard deviation, Variance, IQR, 변동계수(CV) Confusion Matrix Binomial, Bernoulli, Gaussian Distribution 중심극한정리
	통계적 추정의 개념을 이해하고 가설 검정을 수행할 수 있는가	
	이산 및 연속 확률분포를 이해하고 이를 계산 및 응용할 수 있는가	
	데이터 처리를 위한 기초 파이썬 코딩을 할 수 있는가	csv and nested json files Importing and Feature Extraction Feature selection (열 선택, 열 제거, Correlation, Coefficient 분석 등) Missing values, outliers / Time stamp 데이터 정제 noise, inconsistency (데이터 불일치/검증) Standardization, Normalization Discretization, Encoding categorial features Generating new features 필터링 (slicing, indexing, sub-setting) 병합 (merging and joining datasets) 요약 (groupby, aggregation, pivot_table) Line, Bar, Histogram Box, Scatter, Pie Plot Heatmap
	다양한 포맷의 원천 데이터를 데이터프레임으로 로딩할 수 있는가	
	데이터 정제 및 변환 작업을 수행할 수 있는가	
	이종 데이터셋을 결합하고 가공하여, 목적에 맞는 데이터셋으로 가공할 수 있는가	
	주어진 데이터셋과 조건에 적합하도록, Plot을 생성하고 인사이트를 도출할 수 있는가	
	단순, 다중 회귀 개념을 이해하고 회귀분석 모델을 구현할 수 있는가	One variable linear regression Multi variable linear regression Regularization in linear regression Ridge regression / Lasso regression / Data Split K-fold cross validation, LOOCV Overfitting, Underfitting, Bias variance trade-off L1 / L2 regularization Cost Function, Partial derivative Stochastic Gradient Descent / 이항 로지스틱 회귀 Decision boundary (Threshold/Cut-off value) Cost function (Cross entropy loss) / Gradient descent Decision tree learning (Entropy, Information gain, overfitting) Ensemble of decision trees (Random Forest) Regression tree / Classification tree
	대표적 규제 모델인 Ridge, Lasso의 차이점을 구분할 수 있는가	
데이터 처리와 시각화	Train/Validation/Test dataset의 이해 및 활용	
	과적합, 편향-분산 tradeoff 등 문제를 이해하고, 규제 등 대표적 해결 방법들을 구현할 수 있는가	
	가설에 대한 비용 함수를 이해하고, 경사하강법을 구현할 수 있는가	
	로지스틱 모델에 적합한 데이터 특성과 모델의 로짓 함수 등의 개념	
	Decision Tree의 개념과 원리를 이해하고 활용할 수 있는가	
	양상분류의 기본적인 개념과 동작과정을 이해하고 활용할 수 있는가	
	양상분류 모형을 활용하는 배경(이유)과 목적(효과)에 대한 명확한 이해	
	클러스터링 알고리즘(Centroid-based, density-based)의 원리를 이해하고 주어진 데이터에 대한 클러스터링을 수행할 수 있는가	
	SVM의 원리를 이해하고 이를 구현할 수 있는가	
	주요 최적화 목표 함수를 이해하고 올바르게 적용 가능한가	Mean squared error (MSE) Cross Entropy Accuracy Precision, Recall, F1 AUROC, AUPR Feature Importance (+ Visualization) 평가 프로세스 Hyper-parameter의 정의 튜닝 프로세스 Manual Search, Grid Search
ML 모델 빌딩	주요 평가 지표(분류, 회귀, 기타)를 이해, 구현 및 적용 가능한가	
	하이퍼파라미터 튜닝에 대해 이해하고 있는가	
	하이퍼파라미터를 자동으로 최적화하기 위한 방법론에 대해 알고 있는가	
	주요 최적화 목표 함수를 이해하고 올바르게 적용 가능한가	
모델 평가와 최적화	주요 평가 지표(분류, 회귀, 기타)를 이해, 구현 및 적용 가능한가	
	하이퍼파라미터 튜닝에 대해 이해하고 있는가	
	하이퍼파라미터를 자동으로 최적화하기 위한 방법론에 대해 알고 있는가	
	주요 최적화 목표 함수를 이해하고 올바르게 적용 가능한가	

Advanced Practitioner

평가 기준 및 시험 범위는 어떻게 되나요? 1/2



공통 과정

주요 역량	평가 기준	범위
AI Math (DL관련)	이산 및 연속 확률분포를 이해하여, 이를 계산 및 응용할 수 있는가	Discrete / Continuous Distribution Marginal / Conditional Distribution Bayes'Rule Entropy, Conditional Entropy Mutual information, Cross Entropy Kullback-Leibler Divergence Maximum likelihood estimation (MLE) Maximum a posterior (MAP) Naïve BayeLimit, Continuity, Differentiation, Taylor Series Partial differentiation, Gradient Integral, Definite Integral, Indefinite Integral s Classifier eigenvectors, eigenvalues Eigen Decomposition, Spectral Decomposition Singular Value Decomposition
	정보량의 개념을 이해하고, 분포간 엔트로피의 차이를 계산할 수 있는가	
	MLE, MAP의 개념의 차이를 이해하고 활용할 수 있는가	
	다양한 함수를 미분/적분하고, 도함수의 의미를 해석할 수 있는가	
	고유값, 고유벡터 등 개념을 이해하여, 고유값 분해 / 특이값 분해를 수행할 수 있는가	
	딥러닝의 원리와 네트워크 구조를 이해하는가	Perceptron, Forward Propagation Softmax Classifier, Logistic Regression / Backpropagation Pytorch Tensor Operation Forward / Backward Propagation, Autograd Dataset, DataLoader, Custom Layer / Loss Sigmoid, Softmax, Tanh, ReLU Batch Normalization, Layer Normalization Dropout / Data Augmentation Hyper-parameters (Learning rate, batch size, early stopping, weight decay) Hyper-parameter Search (Manual, Grid, Random) Evaluation for Binary Classification (Precision, Recall, F1, AUROC) Evaluation for Binary Classification (Macro-level, Micro-level) Convolution operation / Stacking Convolution Layers Convolutional Neral Networks RNN (Many to One, Many-to-Many) / Language Modeling Exploding & Vanishing Gradient / LSTM, GRU Seq2seq with Attention / Transformer : Scaled-Dot Product Attention, Multi-head Attention, Feed Forward Network, Residual Connection Generative Adversarial Model / Mini-max Optimization Auto-encoder
	딥러닝 프레임워크(Pytorch)를 이해하고, 이를 활용할 수 있는가	
	활성함수(Activation Function)의 종류, 장단점, 특성을 이해하고, 딥러닝 모델에 적용할 수 있는가	
	모델을 학습할 때 발생하는 과적합 문제를 이해하고, 이를 해결할 수 있는 기법을 적용할 수 있는가	
	모델 최적화를 위한 기법을 이해하고 구현할 수 있는가	
DL 모델 빌딩	문제 상황에 맞는 평가지표를 이해하고, 이를 적용할 수 있는가	
	CNN 모델을 이해하고 활용할 수 있는가	
	RNN 모델을 이해하고 활용할 수 있는가	
	Attention에 기반한 모델(Transformer)을 이해 및 활용	
	Generative 모델을 이해하고 활용할 수 있는가	

Advanced Practitioner

평가 기준 및 시험 범위는 어떻게 되나요? 2/2



선택 과정 中 택 1

주요 역량

평가 기준

범위

이미지

이미지 데이터에 대한 전처리 기법을 이해하고 적용할 수 있는가

Image Classification 과제를 이해하고 수행할 수 있는가

Object Detection 과제를 이해하고 수행할 수 있는가

Image Segmentation 과제를 이해하고 수행할 수 있는가

Image Generation 과제를 이해하고 수행할 수 있는가

Vision 분야 주요 과제들을 구현할 수 있는가

Pad, Resize
CenterCrop, FiveCrop
Grayscale, Random transforms
VGGNet
ResNet
MobileNet, Vision Transformer
Two-stage Object Detection (RCNNs)
Evaluation (IoU, NMS)
One-stage (YOLO, SSD)
Upsampling (Sliding Window, unpooling, Transposed conv)
Semantic Segmentation (DeepLab v3)
Instance Segmentation (Masked-RCNN)
Image Generation (DCGAN)
Image Translation (CYCLEGAN)
Image Classification with ResNet
Image Segmentation with Unet
Object Detection with YOLO

자연어

자연어처리의 특성 이해 및 적합한 전처리 과정을 수행할 수 있는가

단어 임베딩과 차원 축소에 대해 이해하였는가

Text Generation 과제를 이해하고 수행할 수 있는가

Pretrained Models의 종류와 학습/활용 방법에 대한 이해

자연어처리 주요 라이브러리(Huggingface)를 활용할 수 있는가

NLP 분야 주요 과제들을 구현할 수 있는가

Regular Expression
English with NLTK (punctuations, stopwords, stemming, lemmatization)
Korean with KoNLPy (morphs, nouns, pos)
Tokenization (Character / Word / Subword basedWord2Vec Implement (CBOW, Skip-gram)
Word2Vec Training and Inference
Word Visualization with Gensim
Decoding (Greedy / Exhaustive / Beam Search)
Evaluation for Text Generation
BERT / GPTs / T5, BART
Huggingface Tokenizer / Huggingface Trainer
Huggingface Model Hubs
Text Classification with LSTM, Transformer (BERT)
Token Classification with LST
Language Modeling with LSTM, GPTM, Transformer (BERT)

시계열

시계열 데이터의 특징 이해 및 적합한 전처리 과정을 수행할 수 있는가

딥러닝 기반의 시계열 예측 모델을 이해하여 구현할 수 있는가

Anomaly Detection 과제를 구현할 수 있는가

ARMA, ARIMA
Decomposition (Trend, Seasonality)
Time Series Forecasting with Transformer
Anomaly Detection (AutoEncoder, anomaly score)

시험 준비를 위한 학습 과정은 mySUNI에서!

mySUNI를 통해 SKADA 인증 획득을 위한 학습이 가능합니다.
본인의 유형에 맞는 학습 과정을 지금 바로 확인해 보세요!



Practitioner

AI Data Analytics
과정을 추천해요! →



학습 과정 바로가기 CLICK!

SKADA 기출 문제를 풀고
문제 유형을 확인해보세요!



시험 대비 문제풀이 CLICK!

Advanced Practitioner

영역별 Adv. AI Data Analytics 과정을 추천해요! →

(*이미지 / 자연어 / 시계열 중 한 개의 영역 선택)

Adv. AI Data Analytics
(이미지)

★★★

[이미지] 과정
CLICK!

Adv. AI Data Analytics
(자연어)

★★★

[자연어] 과정
CLICK!

Adv. AI Data Analytics
(시계열)

★★★

[시계열] 과정
CLICK!

시험 운영 상세 FAQ



Question

Answer

1

시험은 어디에서 보는 건가요?

시험은 오프라인으로 진행되며,
서울권역/이천권역/판교권역 중 선택하실 수 있습니다.

2

시험 신청 Process가 궁금합니다.
비용은 어떻게 되죠?

mySUNI를 통해 신청할 수 있으며,
'24년 인증 비용은 무료로 운영됩니다.

3

Advanced Practitioner의 경우 3개 영역
(자연어, 이미지, 시계열) 모두 합격해야 하나요?

1개 영역만 선택하여 합격하면 인증됩니다.

4

레퍼런스 없이
Code를 작성해야 하나요?

주요 Library의 *Reference site를 즐겨찾기로 제공하며,
그 외의 별도 자료 및 전자기기는 모두 반입 불가합니다.
(*Reference Site 별도 안내)

5

실습 환경은 어떤가요?

Python 언어 기반의 주피터 노트북이 제공되며,
Advanced Practitioner는 Pytorch만 지원합니다.

6

성적은 언제 공개되나요?

4주 이내 개인별 점수가 통보되며,
멤버사 HR에는 Pass/Fail로 공유 됩니다.

7

시험 응시에 횟수 제한이 있나요?

별도의 횟수 제한은 없습니다.

2024년도
SKADA와 함께!
지금 바로 신청하세요.

