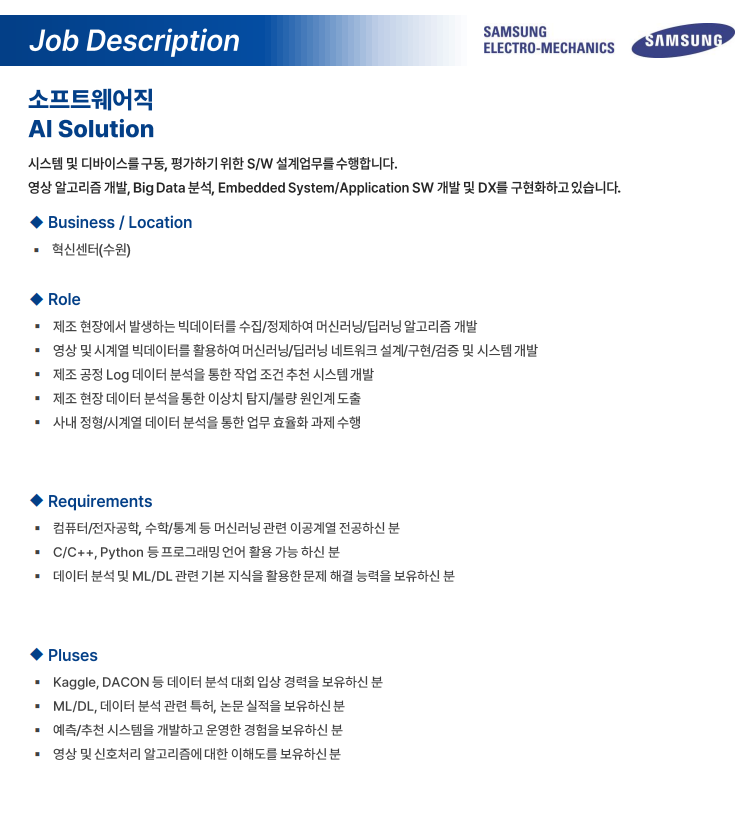
김세훈 삼성전기 AI Solution 최종?



1. **삼성전기를 지원한 이유와 입사 후 회사에서 이루고 싶은 꿈을 기술하십시오. 700자**

[비정형 공정 데이터를 혼합한 분석가]

삼성전기는 자동차의 전자기기화에 힘입어 수요가 증가하고 있는 MLCC와 카메라모듈 등 다양한 전자부품을 생산하고 있기 때문에 성장 가능성을 느껴 지원하게 됐습니다. 그 중 MLCC는 전류수급을 조절하는 제품이기 때문에 불량 검출이 매우 정교해야 합니다. 삼성전기는 불량을 검출하기위해 공정 중 나오는 다양한 데이터를 활용해 AI기반 예측을 진행하고 있습니다. 또한 글로벌 스마트 팩토리 10개 기업에 선정될만큼 정확하고 방대한 데이터를 보유하고 있다는 것이 매력적이라 느꼈습니다.

제가 삼성전기에 입사하게 된다면 정형, 비정형 데이터를 융합한 다양한 데이터를 분석하여 이상치 탐지/불량 원인계를 도출하는 업무를 맡고 싶습니다. 물류센터 자율 주행 로봇을 제작하는 과정에서 팔레트를 인식하기 위해 비정형 데이터를 다뤄봤지만, 정형 데이터와 비정형 데이터를 혼합하여 분석하는 과정은 진행할 수 없었습니다. 예를 들어 공정 중 외관에 일정한 패턴이 발생하며 불량이 탐지된다면 어떤 조건에서 발생하는지 원인을 도출해야 합니다. 이렇게 실무에서는 목표변수가 다양하고 목표변수와 상관계수가 높은 변수가 형식을 가리지 않기 때문에 데이터의 형식에 얽매이지 않는 분석이 필요합니다. 제 데이터 분석능력을 활용해 불량이 발생한다면 원인을 도출해 효율적으로 제품을 생산하며 공정조건을 최적화하는데 일조하겠습니다.

1. **본인의 성장과정을 간략히 기술하되 현재의 자신에게 가장 큰 영향을 끼친 사건, 인물 등을 포함하여 기술하시기 바랍니다. (※작품속 가상인물도 가능) 1500 자**

[영화에서 시작된 빅데이터, AI 분야에 대한 지속적인 관심]

저는 야구를 좋아합니다. 다양한 야구 영화, 드라마를 즐겨 보는 편인데 머니볼이라는 영화를 감명깊게 봤습니다. 야구는 데이터의 스포츠라고 불립니다. 주인공인 빌리 빈은 야구에 데이터 분석을 적용해 제일 눈에 보이는 성과를 낸 인물입니다. 데이터 분석을 활용한 트레이드를 기반으로 2002년 당시 411억이라는 구단 가치로 1565억이었던 뉴욕 양키스와 같은 승수를 기록했고, 현재까지도 유지되는 20연승을 달성했습니다. 데이터는 현재 야구를 비롯한 스포츠를 넘어 다양한 분야에서 의사결정을 위한 도구로 사용하고 있습니다. 저 또한 영화에 감명을 받았고 제가 전공한 전자공학과에서 데이터 분석을 활용하는 방안을 찾고자 했고, 데이터 분석 및 CV분야에 대한 지식을 쌓기 시작했습니다.

이후 공학종합설계를 수강하며 딥러닝 기반 안면 인식 도어락을 제작했고 여러 안면 인식 알고리즘에 대해 학습할 수 있었습니다. Haarcascade 알고리즘과 딥러닝을 활용하는 CNN 알고리즘을 직접 적용해봤습니다. 먼저 Haarcascade 알고리즘은 단순 비교를 활용한 인식이기 때문에 속도는 빠르지만 정확도는 낮은 단점이 있었고, 반면에 CNN 알고리즘은 정확도는 높지만 속도가 느린 단점이 있었습니다. Haarcascade는 100 명 중 30명, CNN은 60명을 인식할 수 있었습니다. 하지만 CNN알고리즘은 라즈베리 파이 내부에서 실시간 인식을 진행할 수 없었기 때문에 Haarcascade로 실시간 CCTV를 구현하고 이후 얼굴이 인식되면 사진을 찍은 후 미리 학습시켜 놓은 얼굴과 대조를 통해 인식을 진행했고 족업작품으로 전시를 진행했습니다. 이 경험을 통해 알고리즘의 중요성을 알게 됐습니다.

또한, AI&Bigdata 분야에 전문성을 갖추고자 AI&Bigdata 교육을 수강했고, 그 과정에서 물류센터 자율 주행 로봇과 카드사 데이터 분석 프로젝트를 진행 했습니다. 카드사 텔레마케팅 데이터를 분석해 사용자의 소비 패턴을 분석했고, 고객의 정보를 입력하면 텔레마케팅 성공률을 도출하는 예측 프로그램을 제작했습니다. 이 과정에서 XGBoost, LightGBM, CatBoost 등 다양한 머신 러닝 모델을 학습했고 적용했습니다. 같은 데이터를 학습시켰는데도 데이터의 성질에 따라 정확도 등 다양한 수치가 변화하는 것을 볼 수 있었습니다. 결과적으로 정확도, 정밀도, F1-score를 고려해 LightGBM모델을 선정했고, 발표한 결과 우수상을 수상했습니다. 이 경험을 통해 데이터에 따라 어떤 알고리즘을 사용해야 하는지 판단하는 능력의 중요성을 깨달을 수 있었습니다.

저는 AI, Bigdata 분야에 지속적인 관심을 가졌고 학습해왔습니다. 제가 학습한 지식을 활용해 삼성전기에서 공정 중 나오는 다양한 데이터를 활용하고 분석하며 고성능의 알고리즘을 제작해 AI Solution 업무를 수행하겠습니다.

1. **최근 사회이슈 중 중요하다고 생각되는 한가지를 선택하고 이에 관한 자신의 견해를 기술해 주시기 바랍니다. 1000자 이내**

[출산율 감소에 따른 대책 - AI, 자동화, Bigdata]

최근 한국은 심각한 출산율 감소 문제를 겪고 있습니다. 출산율을 지속적으로 감소해 22년에는 0.78명대까지 감소했고 결혼 대비 출산율도 감소하는 실정입니다. oecd 평균 출산율은 1.59명으로 평균의 반도 미치지 못하는 수치입니다. 이에 따라 근로 가능 인구는 감소할것이고 지방 및 중소기업은 벌써 근로자가 부족합니다. 특히 물류, 제조업은 기피하는 근로자가 많기 때문에 빠르게 인구 절벽을 대비해야 합니다.

저는 산업의 자동화가 그 해결책이라고 생각합니다. 물류 분야는 완전 자율 주행 자동차, 제조업 분야는 스마트 팩토리가 그 해결책입니다. 현재 다양한 나라들이 공장 자동화를 통한 효율 증가를 진행하고 있습니다. 인더스트리 4.0을 성공적으로 수행해 지능형 생산 시스템을 구축한 독일은 M2M, 빅데이터, 생산 부문과 업무 시스템의 연계를 핵심 기술로 삼고 있으며, 저는 그 중 빅데이터 분야가 가장 중요하다고 생각합니다. 4차 산업 혁명은 최적화가 핵심입니다. 최적화를 하기 위해서는 많은 데이터를 필요로 하기 때문에 데이터를 빠르게, 많이 쌓는 기업이 유리합니다. 글로벌 스마트 팩토리 10개 기업에 선정된 삼성은 다양한 공정 데이터를 기반으로 한국형 스마트 팩토리를 구축하고 그 중 삼성전기는 넥스플랜트(Nexplant)를 기반으로 3.0 버전 MES를 구축했습니다. 지속적으로 데이터를 축적한 삼성은 자동화, 무인화에 누구보다 빠르게 다가갈 수 있을 것입니다.

인구 감소는 빠르게 진행중이고 곧 인구 절벽이 올 것입니다. 인구 감소를 늦추고 향후 다시 인구를 증가시키기 위한 근본적 해결책인 출산율 문제를 해결해야 합니다. 또한 최악의 경우를 대비해 기업 및 국가는 힘을 합쳐 인구 절벽 문제 해결책인 자동화, 무인화의 기술 발전과 제도 마련이 시급하다고 생각합니다.

1. **지원을 희망하는 직무와 관심있는 제품에 대해 작성하고, 전공지식/경험 (프로젝트, 논문, 공모전 등)을 토대로 본인이 적합한 사유를 서술하시오. 1000자**

[다양한 데이터를 쌓아야 하기 때문에 빅데이터 분석 역량이 필수적이다.]

AI 솔루션 직무를 희망하고 MLCC에 대해 관심을 가지고 있습니다. 현재 삼성전기는 AI를 활용해 MLCC의 품질을 판단합니다. 전자기기는 점점 작아지고 있으며 부품의 크기도 감소하고 있습니다. 그래서 육안으로는 품질검사를 실시할 수 없기 때문에 정확한 AI를 활용한 검사가 필수입니다. AI 솔루션 직무는 공정 중 생성되는 다양한 데이터를 다루기 때문에 데이터 분석 역량이 중요합니다. 저는 데이터 상관관계 분석이 중요하다고 생각합니다. 현재 문제가 발생하거나, 예측이 필요한 데이터와 그 데이터를 예측하기 위한 데이터를 골라내야 효율적인 분석이 가능하기 때문입니다.

저는 미세먼지 데이터 분석 프로젝트를 진행하며 데이터의 상관관계를 통해 결측치를 예측한 경험이 있습니다. 데이터를 수령했을 때 결측치가 많은 것을 발견했습니다. 일산화탄소 데이터가 한 달 가량 결측돼있었고, 대체할 방법을 찾아야 했습니다. 논문을 참조한 결과 일산화탄소는 미세먼지 발생에 있어 중요한 요소여서 기존의 평균값, 중앙값을 이용한 방식은 정확도가 떨어질 것이라는 생각이 들었습니다. 정확한 예측값을 찾기 위해 선형회귀를 이용하기 위해 히트맵을 생성해 일산화탄소와 상관계수가 높은 이산화질소를 찾았습니다. 이후 이산화질소를 활용해 선형 회귀로 일산화탄소를 예측했습니다. 이후 일산화 탄소, 아황산가스가 미세먼지 발생에 큰 영향이 있는 것을 알 수 있었습니다. 두 유해물질을 경감하기 위한 대책을 조사한 후 발표를 진행했고, 교수님에게 좋은 피드백을 들을 수 있었습니다.

공정 데이터는 다양한 센서를 활용하기 때문에 센서가 고장나거나 오류가 있다면 이상치, 결측치가 생성되기 때문에 이상치, 결측치에 대한 처리가 필수입니다. 제가 진행한 프로젝트를 기반으로 더 정확한 데이터 예측 프로그램을 만드는데 기여하겠습니다.