

생물다양성 모델링과 시뮬레이션 및 예측 알고리즘

이정민/컴퓨터학과/2023320060/고려대학교



개요

I. 서론

1. 생물다양성 모델링 및 시뮬레이션 소개
2. 예측 알고리즘 소개

II. 생물 다양성 모델링과 시뮬레이션, 알고리즘의 활용

III. 활용으로 얻을 수 있는 이점

IV. 한계점

V. 결론

I. 서론

1. 생물 다양성 모델링 및 시뮬레이션이 무엇인가?

생물 다양성 모델링은 생물학적인 시스템(계)에서 다양한 생물 종 간의 상호작용과 생태계의 복잡한 구조를 이해하고 예측하기 위한 모델링 기술을 가리킵니다. 모델링 기술을 적용하는 과정에서는 많은 환경 변수와 물리적/비물리적 요소를 고려하는 등 현실의 생태계와 유사하게 구현합니다. 각각의 객체가 가지고 있는 행동과 상호작용을 모방하여 전체 시스템의 흐름을 모델링하는 에이전트 기반 모델링이 예시가 될 수 있습니다. 컴퓨터 과학에서 모델링은 주로 수학적 모델, 컴퓨터 시뮬레이션, 통계적 모델, 그래프 및 네트워크 모델 등을 사용하여 이루어집니다. 오늘 저희가 살펴볼 생물다양성 모델링은 생태학, 생물학, 컴퓨터 과학, 통계학 등 다양한 분야의 지식을 통합하여 사용됩니다. 이를 통해 생태계의 복잡한 상호작용을 이해하고 보전 및 관리에 도움이 되는 정보를 제공합니다. 모델링 된 프로그램을 시뮬레이션을 통해 구동하여 일종의 가상 생태계를 구성할 수 있습니다. 좀 더 쉽게 말하자면, 전산 세계 내부에 하나의 생태계를 구현하는 것을 시뮬레이션이라고 말할 수 있습니다.

2. 예측 알고리즘 소개

생물 다양성의 모델링을 위해 사용되는 알고리즘에는 유전 알고리즘, 기계 학습 알고리즘 등 많은 알고리즘이 존재하지만 이번 과제에서는 유전 알고리즘과 기계 학습 알고리즘에 관해서만 살펴볼 것입니다. 유전 알고리즘은 생물학적 진화 개념에서 영감을 받아 문제 해결을 위한 최적화 기술로, 생물다양성 내의 진화 및 유전적 다양성 연구에 적용됩니다. 예를 들어, 어떤 종의 특정 유전자가 생존에 어떤 영향을 미치는지 이해하는 데 사용될 수 있습니다. 기계 학습 알고리즘은 데이터를 기반으로 패턴을 학습하고 예측하는 데 사용됩니다. 이는 생물다양성 데이터의 분류, 멸종 위험 평가 등에 활용될 수 있습니다. 예를 들어, 생물학적 데이터를 통해 종 다양성의 변화를 예측하거나, 특정 환경 조건에서 특정 생물종의 생존 가능성을 평가하는 데 사용될 수 있습니다.

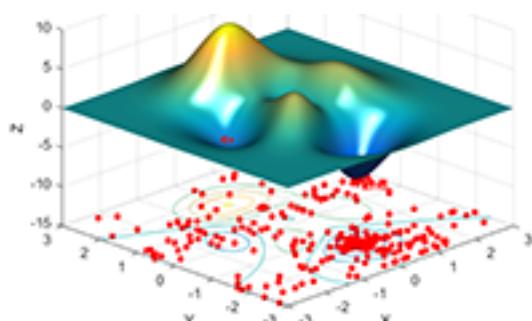


그림 1. 유전 알고리즘의 적용 사진/ 출처 : matlab

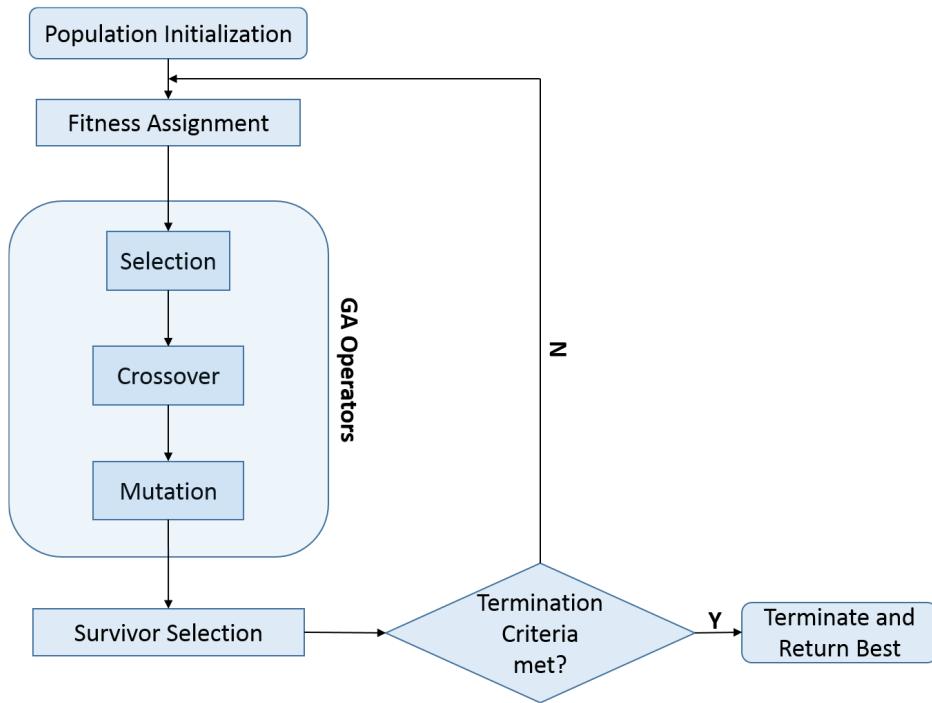


그림 2. 유전 알고리즘의 6단계 / 출처 : <https://github.com/pilsung-kang/Business-Analytics-IME654->

II. 생물 다양성 모델링과 시뮬레이션, 알고리즘의 활용

저희는 생물 다양성 모델링과 알고리즘을 적용하여 다음의 목적들을 달성할 수 있습니다.

1. 종 멸종 위험 예측

생물 다양성 모델링을 사용한다면, 특정 생물종의 멸종 위험을 예측할 수 있습니다. 이를 통해 어떤 종이 위기에 처해 있으며 어떤 요인들이 이러한 위험을 증가시키는지를 파악하여 보전 활동을 계획할 수 있습니다.

2. 환경 변화에 대한 생물 종 반응 예측

생물 다양성 모델링을 사용하여 기후 변화, 환경 오염 등과 같은 환경 요인이 생물종에 미치는 영향을 예측할 수 있습니다. 이를 통해 어떤 환경 변화가 생물 종 다양성에 영향을 미치는지 이해하고, 보전 정책을 개발할 수 있습니다.

3. 보전 지역 선정과 자원 할당:

모델을 사용하여 생물다양성이 풍부한 지역을 식별하고, 어떤 지역에서 어떤 생물종이 가장 취약한지를 파악하여 보전 자원을 효율적으로 할당할 수 있습니다.

4. 보전 계획 평가 및 실행

모델링은 보전 계획을 평가하고 실행하기 전에 시뮬레이션을 통해 효과를 예측하는 데 사용될 수 있습니다. 이를 통해 어떤 보전 전략이 가장 효과적일지를 예측하여 채택할 수 있습니다.

생물다양성 모델링과 알고리즘을 사용하여 생태계의 상호작용과 생물 종 다양성을 이해하고 예측하여 보전 및 관리에 도움을 줄 수 있습니다. 이와 관련하여 독자적인 알고리즘을 개발하여, 생물 다양성 데이터 분석을 통해 유의미한 결과를 가져온 사례도 많은데, 대표적인 예시에는 2022년 9월에 열린 생물다양성 데이터분석 및 아이디어제안 경연(AI Challenge for Biodiversity, 이하 생물다양성 경연)의 데이터분석 리그에서 수상한 sticksINbundle 팀의 시민과학 데이터를 활용한 곤충 종의 연간 분포 패턴 및 감소율 예측입니다. 북미 토착 무당벌레 4종의 연간 감소율을 예측하는 머신러닝(ML) 모델을 개발하였습니다. 이 머신러닝 모델이 특이한 점은 분석 시 주로 쓰이는 환경변수(온도 · 고도 등)는 제외하고, 생물변수(주변 생물)만 활용하는 등 독자적인 아이디어도 접목했다는 점입니다. 환경변수를 활용할 경우 수십 년의 기간을 놓고 봐야 하는데, 생물변수를 통해 더 짧고 세밀한 기간 내 생물종 서식 여부를 예측할 수 있게 됐습니다. 그 결과 4종의 무당벌레 모두 매해 감소가 예상되며, 특히 3개 종은 '멸종위험군'으로 분류된다는 결론을 얻었습니다. 북미 토착 무당벌레 4종은 멸종위기종 등급 DD에 해당하기 때문에 해당 연구는 매우 높은 가치를 가지고 있습니다. 이처럼 생물 다양성 모델링과 그 알고리즘은 그 자체로도, 혹은 응용되어 생물다양성 분야에 넓게 자리잡아 기여하고 있습니다.

III. 활용으로 얻을 수 있는 이점

생물 다양성 모델링과 시뮬레이션, 알고리즘을 활용하여 얻을 수 있는 이점으로는 보전 및 관리에 대해 효율성이 증대된다는 점이 있습니다. 우리는 모델링을 통해 생태계의 복잡한 상호작용을 이해하고 예측할 수 있어서, 자연 생태계의 보전 및 관리에 효율적인 전략을 개발할 수 있어서 보전 노력의 효과를 예측하고 최적화하여 자원을 최대한 효율적으로 활용할 수 있게 됩니다.

또한 시뮬레이션을 통해 이루어지기 때문에 물리적인 시간이 비교적으로 적게 걸린다는 점도 큰 이점 중 하나입니다. 현실의 생태계는 매우 복잡하고 유기적인 흐름을 지니고 있기 때문에 한 종에 변화가 일어나기까지 오랜 시간이 걸립니다. 시뮬레이션에서는 전혀 관련 없는 일부 변수들은 배제하거나, 혹은 가속화하여 진행할 수 있기 때문에 비교적 짧은 시간에 유의미한 결과를 도출 할 수 있습니다.

마지막으로는 실제 생태계의 파괴가 일어나지 않는다는 점입니다. 지금까지의 인류는 우리의 삶의 터전이 파괴되는 것을 몸소 체험하고 경험하며 그 경각심을 일깨워왔습니다. 하지만 경각심을 지닌 지금의 시기는 매우 늦은 타이밍으로 회복이 불가능할 정도로 진행된 부분도 분명히 존재합니다. 모델링을 적용한다면, 가상의 생태계가 파괴되는 것으로 미리 예측하여, 우리의 삶의 터전을 지킬 수 있습니다.

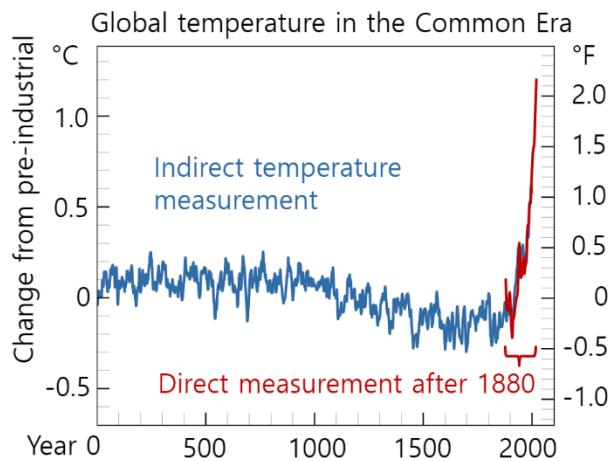


그림 3. 평균 기온의 상승 그래프

IV. 한계점

생물 다양성 모델링과 알고리즘은 많은 이점을 가지고 있지만 몇 가지 한계점도 고려해야 합니다. 먼저 데이터의 부족 및 불충분성으로 인한 정확도 하락이 있습니다. 생물 다양성 모델링에 필요한 데이터의 부족이나 불충분성은 모델의 정확도를 제한할 수 있습니다. 특히, 특정 지역이나 생물종에 대한 충분한 데이터가 없는 경우 모델의 신뢰성이 저하될 수 있습니다. 또한 생태학적 시스템이 매우 복잡하고, 수많은 상호작용으로 이루어져 있다는 점입니다. 실제 생태계의 모든 변수를 고려하기 어려워 예측이나 모델의 정확도가 한계를 가질 수 있습니다. 가장 현실적인 문제로는 기술적 한계가 있습니다. 복잡한 모델링과 시뮬레이션은 계산 리소스와 컴퓨팅 파워를 많이 필요로 하며, 이는 비용이나 기술적 제한으로 이어질 수 있습니다. 이러한 한계점들은 모델링과 알고리즘을 사용할 때 고려해야 하는 중요한 측면들입니다. 이를 고려하여 모델링을 신중하게 수행하고, 모델의 결과를 평가하고 해석함으로써 모델의 한계를 극복하고 향후 보완해야 할 것입니다.

V. 결론

생물 다양성과 모델링 그리고 이와 관련된 알고리즘에 대해서 지금까지 소개해드렸습니다. 컴퓨터 과학 분야와 생물다양성은 전혀 관련 없는 분야라고 생각했지만, 이번 조사를 통해 매우 큰 연결점이 있음을 알 수 있었습니다. 지구의 생물다양성은 인류가 해결해나가야 할 중요한 문제들 중 하나라는 점에서 앞으로 더 많은 기술이 개발되어야 할 것입니다.

출처

유전 알고리즘

<https://github.com/pilsung-kang/Business-Analytics-IME654->

https://yngie-c.github.io/machine%20learning/2020/09/07/genetic_algo/

https://ko.wikipedia.org/wiki/%EC%9C%A0%EC%A0%84_%EC%95%8C%EA%B3%A0%EB%A6%AC%EC%A6%98

기계 학습

https://ko.wikipedia.org/wiki/%EA%B8%B0%EA%B3%84_%ED%95%99%EC%8A%B5

생물 다양성 모델링

<https://news.skhynix.co.kr/post/ai-challenge-for-biodiversity>

<https://www.dbpedia.org/pdf/pdfView.do?nodeId=NODE11418085> / 생태계 및 생물다양성 보전을 위한 종분포 모델/ 박영석/ 경희대학교 생물학과 교수