

Användning av Bayesianisk modellering och bildigenkänning för att kartlägga invasiva växtarter.

Arbetar du själv eller med någon:

Jag arbetar självständigt med detta projekt.

Kort sammanfattning:

Projektet undersöker hur Bayesian-modellering och bildigenkänning kan samverka för att kartlägga invasiva växtarter i Västra Götalands län, Sverige. Genom att integrera data från olika källor, utvecklar jag en metod för att identifiera och kartlägga dessa arter och deras spridning. Bildigenkänning används för att extrahera viktiga egenskaper från bilder, vilka sedan integreras i Bayesian-modellen för att förbättra prestanda och noggrannhet. Fusion av resultat från båda modellerna ger en mer omfattande bedömning av arterens förekomst och ekologiska mönster, medan Bayesian-modellens osäkerhetsuppskattning bidrar till pålitliga prediktioner. Slutsatserna valideras genom att jämföra dem med bildigenkännings resultat, vilket möjliggör en bedömning av modellens tillförlitlighet och identifiering av förbättringsområden.

Syfte och frågeställningar:

Syftet med uppsatsen är att undersöka hur Bayesianisk modellering och bildigenkänning kan användas för att effektivt identifiera och kartlägga invasiva växtarter i Västra Götalands län. Den ökande globala rörligheten för människor och handel har ökat risken för introduktion och spridning av invasiva växtarter, vilket utgör allvarliga hot mot ekosystemens stabilitet och den biologiska mångfalden. Trots betydelsen av att bekämpa invasiva arter är kartläggning och övervakning av dem ofta en utmaning på grund av bristen på tillförlitliga och effektiva metoder.

Denna forskning är av stor betydelse för bevarandet av den biologiska mångfalden och för att möjliggöra effektiva åtgärder för att hantera invasiva arter. Den stöds av en pågående informationssatsning från Naturvårdsverket och andra myndigheter och organisationer för att öka allmänhetens medvetenhet om invasiva arter och deras skadliga effekter på miljön. Genom att utveckla och validera metoder för att kartlägga invasiva växtarter kan denna studie bidra till att öka förståelsen för deras förekomst och spridning, vilket i sin tur kan leda till effektivare strategier för deras kontroll och förvaltning. Dessutom har en modell som kan identifiera växter på bilder potential att bli ett användbart verktyg för allmänheten, vilket kan bidra till att öka medvetenheten om invasiva arter och underlätta deras tidiga upptäckt och hantering.

Forskningsfrågor:

- 1 Vilken roll kan Bayesianisk modellering och bildigenkänning spela i att upptäcka och kartlägga invasiva växtarter i Västra Götalands län?

Genom att integrera data från olika källor och använda avancerade algoritmer för bildanalys och modellering, undersöker denna forskning hur Bayesianisk modellering och bildigenkänning kan bidra till att identifiera och kartlägga invasiva växtarter i regionen.

- 2 Hur kan olika datakällor integreras för att förbättra noggrannheten och tillförlitligheten i identifieringen av invasiva arter?

Detta syftar till att utforska metoder för att integrera data från olika källor, inklusive myndigheters databaser, vetenskapliga artiklar och bildobservationer från medborgare, för att förbättra precisionen och tillförlitligheten i identifieringen av invasiva växtarter.

Utmaningar:

1 Ett potentiellt hinder i detta projekt är att integrera data från olika källor och säkerställa deras tillförlitlighet och relevans för modelleringen av invasiva arter, vilket kan vara tidskrävande.

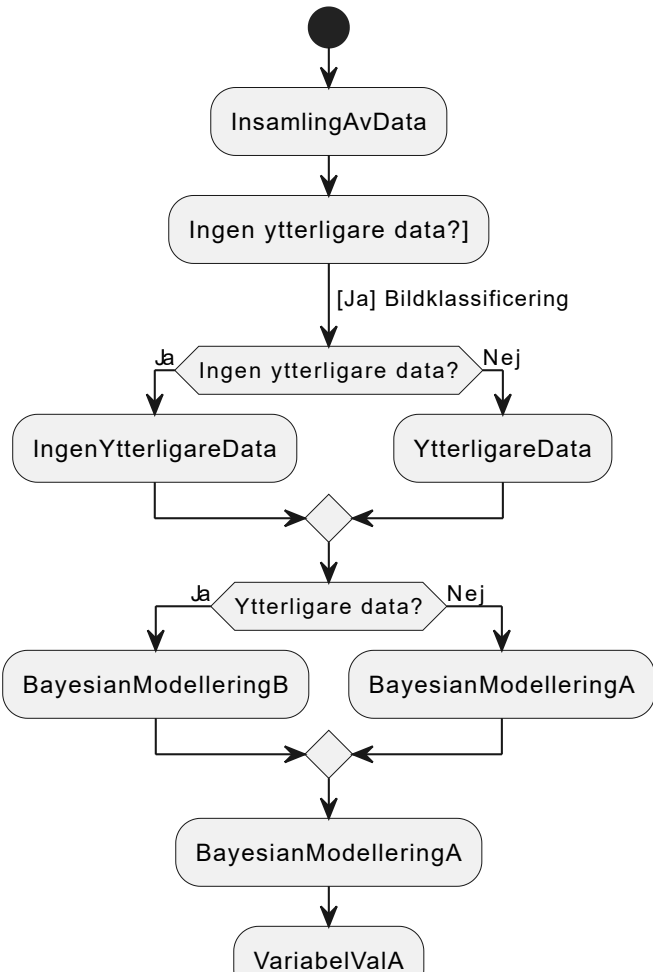
2 Dessutom kan det ta tid att finjustera modellen och hitta optimala hyperparametrar.

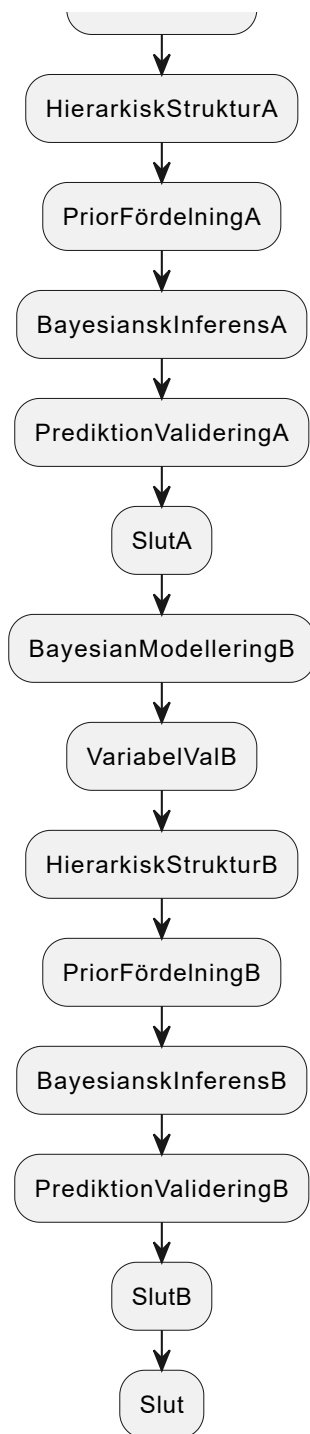
- En möjlig lösning är att använda en bas-modell som är tränad på växter generellt som transfer-learning, men inte specifikt på invasiva arter.

3 Bildigenkänning kan vara komplicerat på grund av variationer i bildkvalitet, ljusförhållanden och vinkel, vilket kan leda till osäkerhet i identifieringen av växtarter.

4 Bayesianisk modellering involvera komplexa statistiska beräkningar och antaganden, vilket kan leda till osäkerhet i modellens prediktioner och resultat om inte jag klarar av at göra dom rätt eller gör fel antagenden.

Utkast av plan





Kod exempel för steg A till E.

```
with pm.Model() as bayesian_model:
    # A) Variabelval
    # Priors för logistiska regressionskoefficienter
    # Tilldela en normalpriorfördelning till koefficienterna (beta) med
    # medelvärde 0 och standardavvikelse 10
    beta = pm.Normal('beta', mu=0, sd=10, shape=X_train_scaled.shape[1])

    # B) Skapa hierarkisk struktur
    # Example: location
    # Prior för medelvärdet av normalfördelningen (mu) för varje plats
```

```
mu_plats = pm.Normal('mu_plats', mu=0, sd=10, shape=antal_platser)

# Prior för standardavvikelsen för normalfördelningen (sd) för varje plats
sd_plats = pm.HalfNormal('sd_plats', sd=1, shape=antal_platser)

# Prior för plats-specifik effekt
beta_plats = pm.Normal('beta_plats', mu=mu_plats, sd=sd_plats,
shape=antal_platser)

# C) Tilldela priorfördelningar
# Omdefiniera priorfördelningen för koefficienterna (beta)
beta = pm.Normal('beta', mu=0, sd=5, shape=X_train_scaled.shape[1]) #
Exempel på ändring av prior

# D) Bayesiansk inferens
# Beräkna sannolikheten för att målbladet ska vara 1 med hjälp av en
logistisk funktion applicerad på punktproduktet av funktioner och koefficienter
p = pm.math.sigmoid(pm.math.dot(X_train_scaled, beta) +
beta_plats[plats_index])

# E) Prediktion och validering
# Definiera Bernoulli-likheten med observerade målblad
# Använd en Bernoulli-likhet för att modellera binära resultat (0 eller 1)
med sannolikheter som ges av p
likelihood = pm.Bernoulli('y', p=p, observed=y_train)
```

Källor:

- 1 Lista med arter som är invasiva. <https://rapportera.artfakta.se/eftersokta/ias/taxa?swe=1>
- 2 Länsstyrelsen Västra Götaland, Invasiva främmande arter <https://www.lansstyrelsen.se/vastra-gotaland/djur/invasiva-frammande-arter.html>
- 3 NATURVÅRDSVERKET, Definition <https://www.naturvardsverket.se/amnesomraden/invasiva-frammande-arter/vad-ar-ifa/definition/>
- 4 Invasiva främmande arter på EU:s förteckning – Växter <https://www.naturvardsverket.se/4acbae/globalassets/amnen/invasiva-frammande-arter/pdf/forteckning-av-invasiva-vaxter/eu-listade-invasiva-frammande-arter-vaxter.pdf>

Tillgång till data:

- 1 Bildobservationer av växter från iNaturalist. Innehåller många datapunkter (time, position, taxonomy etc.) <https://www.inaturalist.org/observations/export>
- 2 Artdatabanken och Artportalen <https://www.artdatabanken.se/tjanster-och-miljodata/artportalen/>
<https://www.artdatabanken.se/>
 - hundratals bilder på växter.
 - en karta med fynd som man kan exportera.

3 TRY, en centraliserad databas för växtträdata och integrerar hundratals dataset- innehållande mer än 300 000 växttaxa och 2 661 egenskaper. <https://www.try-db.org/de/de.php>

Tidsplanering:

Jag planerar att följa nedanstående tidsram för mitt projekt:

Vecka 1: Ladda ner data och samla in det nödvändiga materialet. Städa och bearbeta data. Skriva teoriavsnittet som kommer att användas och slutföra bearbetningen av data.

Vecka 2: Genomföra explorativ dataanalys (EDA). Påbörja AI-modelleringen.

Vecka 3: Fortsätta med AI-modelleringen och genomföra Bayesian analys. Sträva efter att skapa en effektiv AI-modell.

Vecka 4: Skapa visuellt tilltalande diagram och plots. Färdigställa rapporten och genomföra en noggrann korrekturläsning. Förbereda en presentation för opponeringen.