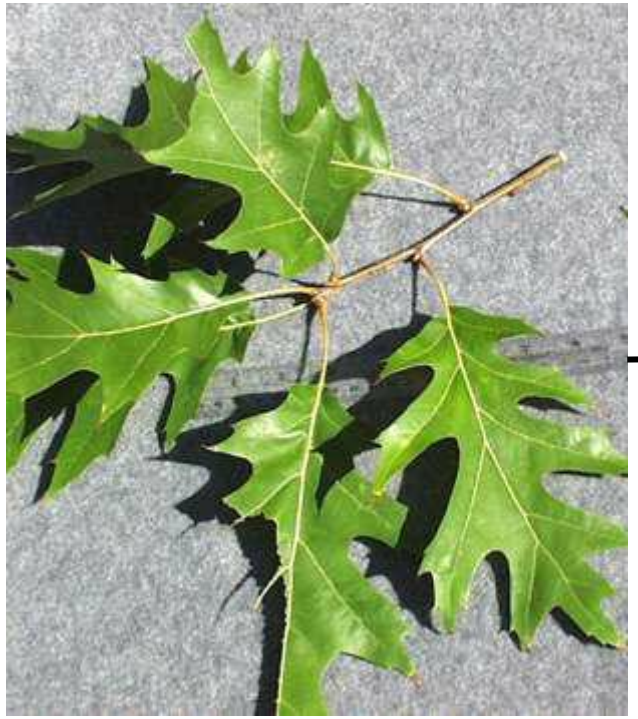
Automatic plant trait annotation for plant species ID

Diego Marcos and Robert van de Vlasakker

**Robert van de Vlasakker,
Wageningen University**

Intro | A classical CNN



Black Box

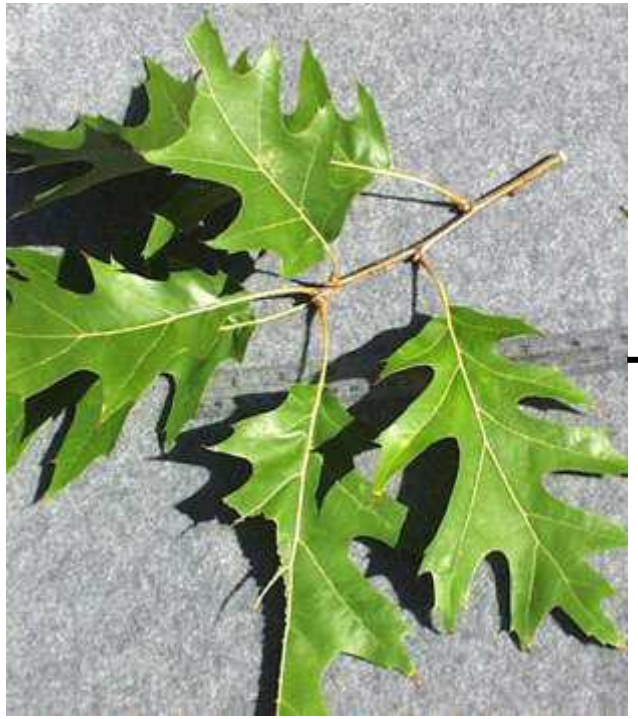
How to handle new species?

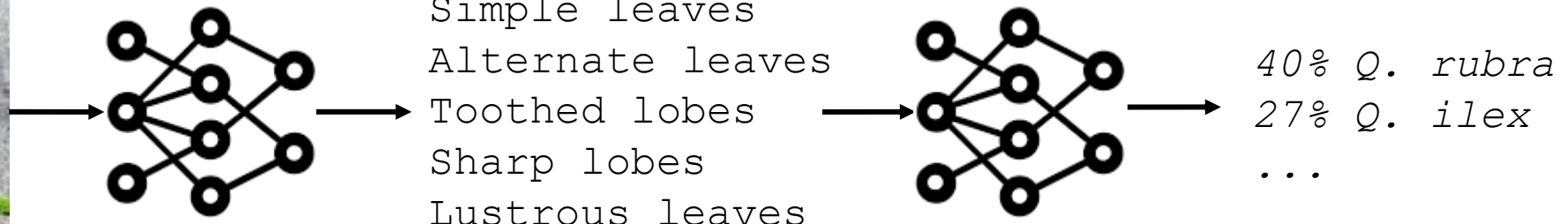
40% Quercus rubra
27% Quercus ilex
...

Intro | Visual-Language Hybrid Model

+ Interpretability

+ Zero-shot and few-shot learning





Preliminary work on this:

Semantically Interpretable Activation Maps: what-where-how explanations within CNNs. D Marcos, S Lobry, D Tuia. ICCVW 2019

Contextual Semantic Interpretability. D Marcos, R Fong, S Lobry, R Flamary, N Courty, D Tuia. ACCV 2020

Where to get data | Morphological trait databases

Existing trait databases tend to be limited in scope in terms of **geography** or **taxonomic group** and focus more on **physiological** than **morphological** traits.

nature › scientific data › data descriptors › article

Data Descriptor | Open Access | Published: 30 September 2021

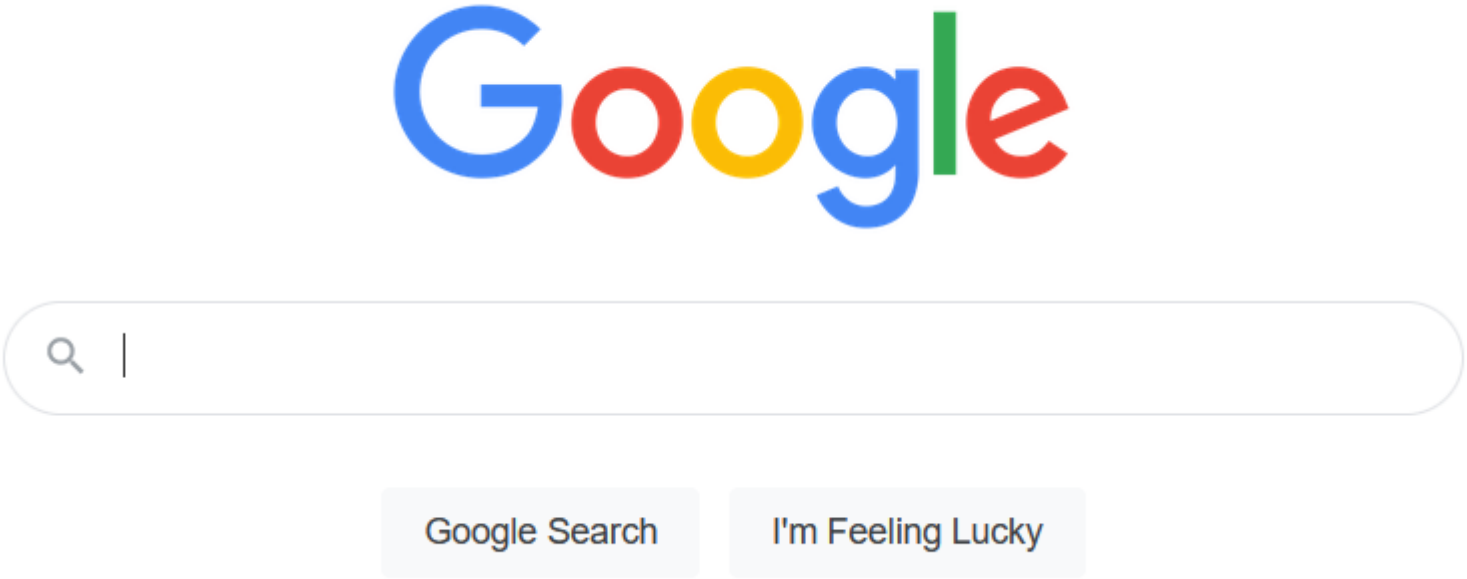
AusTraits, a curated plant trait database for the Australian flora

nature › scientific data › data descriptors › article

Data Descriptor | Open Access | Published: 24 September 2019

PalmTraits 1.0, a species-level functional trait database of palms worldwide

Where to get data | The World Wide Web



Google

quercus rubra description

All Images Shopping News Maps More Tools

About 1,720,000 results (0.55 seconds)

Quercus rubra, commonly called red oak or northern red oak, is **a medium sized, deciduous tree with a rounded to broad-spreading, often irregular crown**. Typically grows at a moderate-to-fast rate to a height of 50-75' (often larger in the wild).

Suggested Use: Shade Tree, Street Tree

Common Name: red oak

Type: Tree

Native Range: Eastern North America

https://www.missouribotanicalgarden.org › PlantFinder

Quercus rubra - Plant Finder - Missouri Botanical Garden

About featured snippets • Feedback

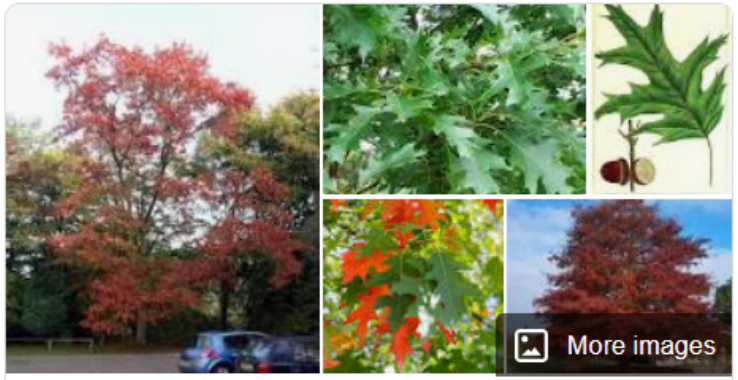
https://en.wikipedia.org › wiki › Quercus_rubra

Quercus rubra - Wikipedia

Description — DescriptionEdit · **Bark: Dark reddish gray brown, with broad, thin, rounded ridges, scaly**. · Wood: Pale reddish brown, sapwood darker, heavy, ...

Section: Quercus sect. Lobatae
Family: Fagaceae
Species: Q. rubra
Kingdom: Plantae

Description · Distribution and habitat · Ecology · Uses



Northern Red Oak

Plant

Quercus rubra, the northern red oak, is an oak tree in the red oak group. It is a native of North America, in the eastern and central United States and southeast and south-central Canada. It has been introduced to small areas in Western Europe, where it can frequently be seen cultivated in gardens and parks. Wikipedia

Scientific name: Quercus rubra

Family: Fagaceae

Kingdom: Plantae

Order: Fagales

Where to get data | The World Wide Web

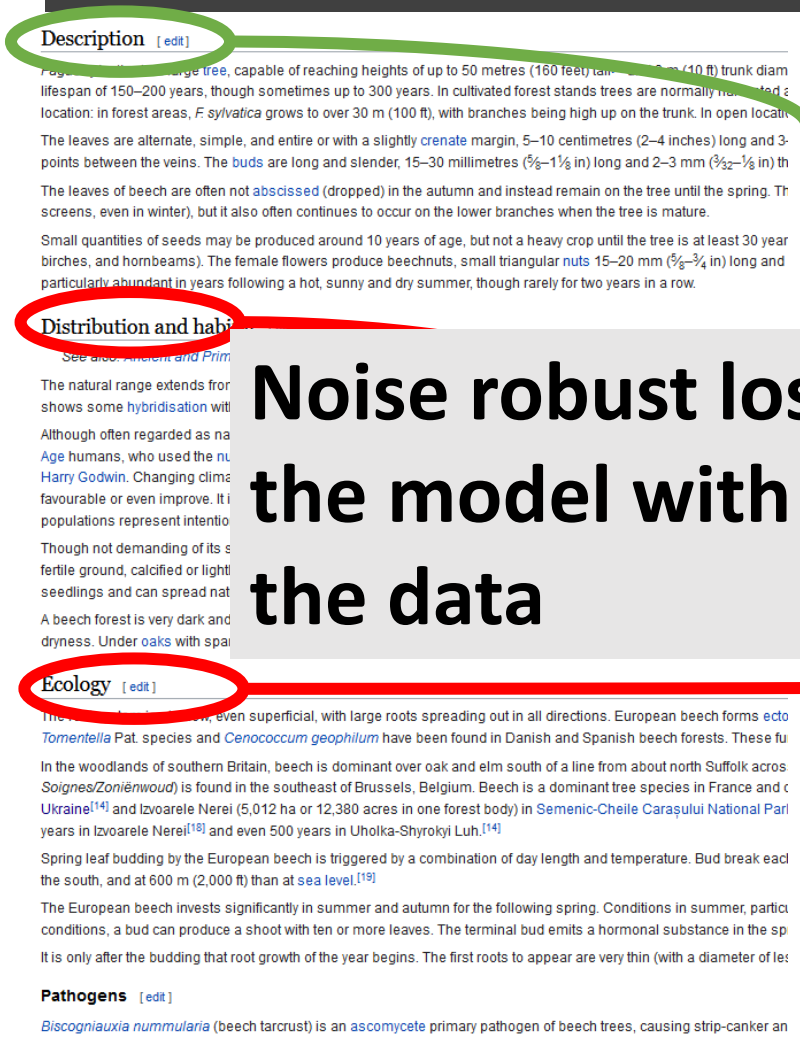
Description paragraph from the first page in Google

This 75-100 ft., deciduous oak occasionally reaches 120 ft. in height. Its straight trunk is clear of branches for some distance above the ground and supports a wide canopy, commonly 3/4 that of height. The dark bark is striped with long, smooth plates separated by deep furrows. Leaf lobes are bristle-tipped. Fall color is can be crimson, golden-orange, or russet.

Bristle-tipped leaves turn red in the fall. The leaves have 7 to 11 waxy lobes. A good street tree, tolerates pollution and compacted soil. Grows as much as two feet a year for 10 years. Grows to 60' to 75', 45' spread. (zones 3-8)

Northern red oak (*Quercus rubra*) is a medium-sized deciduous tree native to the eastern United States, where it has a wide distribution range. The tree was introduced to Europe in the 18th and 19th centuries and is now naturally found throughout western and central Europe.

Where to get data | Binary Description Classifier



Description [edit]

...large tree, capable of reaching heights of up to 50 metres (160 feet) tall ... (10 ft) trunk diam... lifespan of 150–200 years, though sometimes up to 300 years. In cultivated forest stands trees are normally ha...ted a... location: in forest areas, *F sylvatica* grows to over 30 m (100 ft), with branches being high up on the trunk. In open locati...

The leaves are alternate, simple, and entire or with a slightly crenate margin, 5–10 centimetres (2–4 inches) long and 3-... points between the veins. The buds are long and slender, 15–30 millimetres (⅝–1⅛ in) long and 2–3 mm (3⁄32–⅛ in) th...

The leaves of beech are often not abscissed (dropped) in the autumn and instead remain on the tree until the spring. Th... screens, even in winter), but it also often continues to occur on the lower branches when the tree is mature.

Small quantities of seeds may be produced around 10 years of age, but not a heavy crop until the tree is at least 30 year... birches, and hornbeams). The female flowers produce beechnuts, small triangular nuts 15–20 mm (⅝–¾ in) long and ... particularly abundant in years following a hot, sunny and dry summer, though rarely for two years in a row.

Distribution and habi...

See also: Ancient and Prim...

The natural range extends fror... shows some hybridisation witl...

Although often regarded as na... Age humans, who used the nu... Harry Godwin. Changing clima... favourable or even improve. It i... populations represent intentio...

Though not demanding of its s... fertile ground, calcified or lightl... seedlings and can spread nat...

A beech forest is very dark and... dryness. Under oaks with spa...

Ecology [edit]

...even superficial, with large roots spreading out in all directions. European beech forms ecto... *Tomentella* Pat. species and *Cenococcum geophilum* have been found in Danish and Spanish beech forests. These fu...

In the woodlands of southern Britain, beech is dominant over oak and elm south of a line from about north Suffolk acros... *Soignes/Zoniënwoud*) is found in the southeast of Brussels, Belgium. Beech is a dominant tree species in France and ... Ukraine[14] and Izvoarele Nerei (5,012 ha or 12,380 acres in one forest body) in Semenic-Cheile Carașului National Parl... years in Izvoarele Nerei[18] and even 500 years in Uholka-Shyrokyi Luh.[14]

Spring leaf budding by the European beech is triggered by a combination of day length and temperature. Bud break eacl... the south, and at 600 m (2,000 ft) than at sea level.[19]

The European beech invests significantly in summer and autumn for the following spring. Conditions in summer, particu... conditions, a bud can produce a shoot with ten or more leaves. The terminal bud emits a hormonal substance in the sp...

It is only after the budding that root growth of the year begins. The first roots to appear are very thin (with a diameter of les...

Pathogens [edit]

Biscogniauxia nummularia (beech tarcrust) is an ascomycete primary pathogen of beech trees, causing strip-canker an...

Noise robust loss that helps the model with the noise in the data

Used as ... The text DOES contain description ...

Actually not completely True!

Used as Negative labels: The te... contain descripti...

Actually not completely True!

Where to get data | The World Wide Web

Trained on 3 structured data sources:

- Wikipedia (pages about plants, birds and mammals, enriched with non-organism pages)
- Birds of the World (pages about birds)
- Plants of the World Online (pages about plants)

Tested on 2 completely left out data sources:

- Llifle (pages about plants)
- AgroForestry (pages about plants)

Where to get data | The World Wide Web

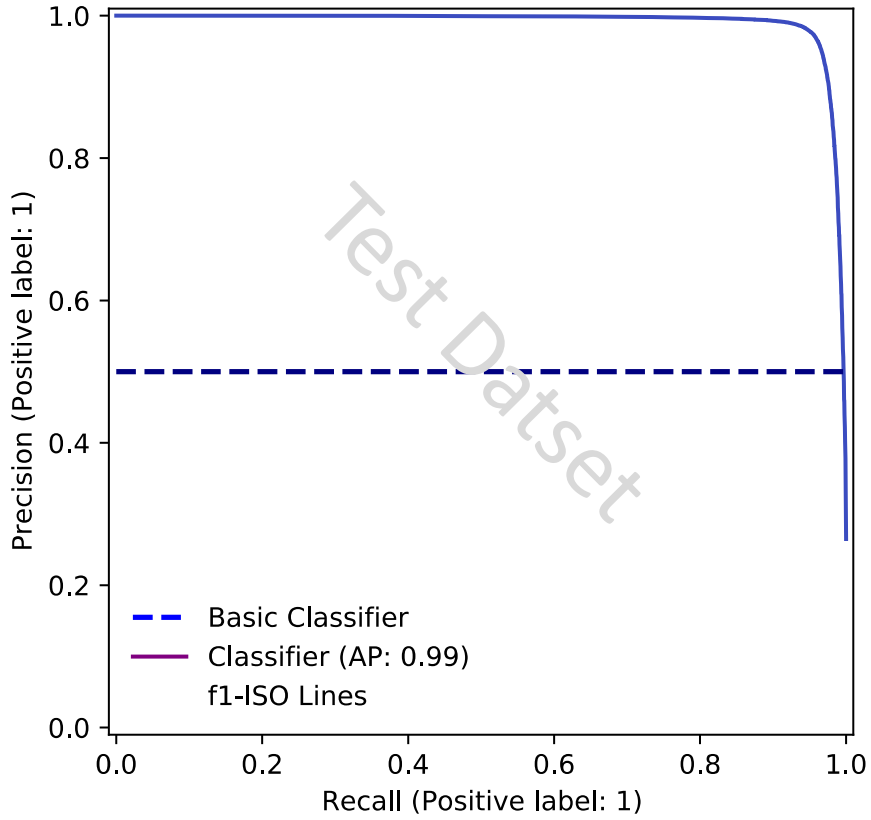
Over 50,000 species and pages used for training:

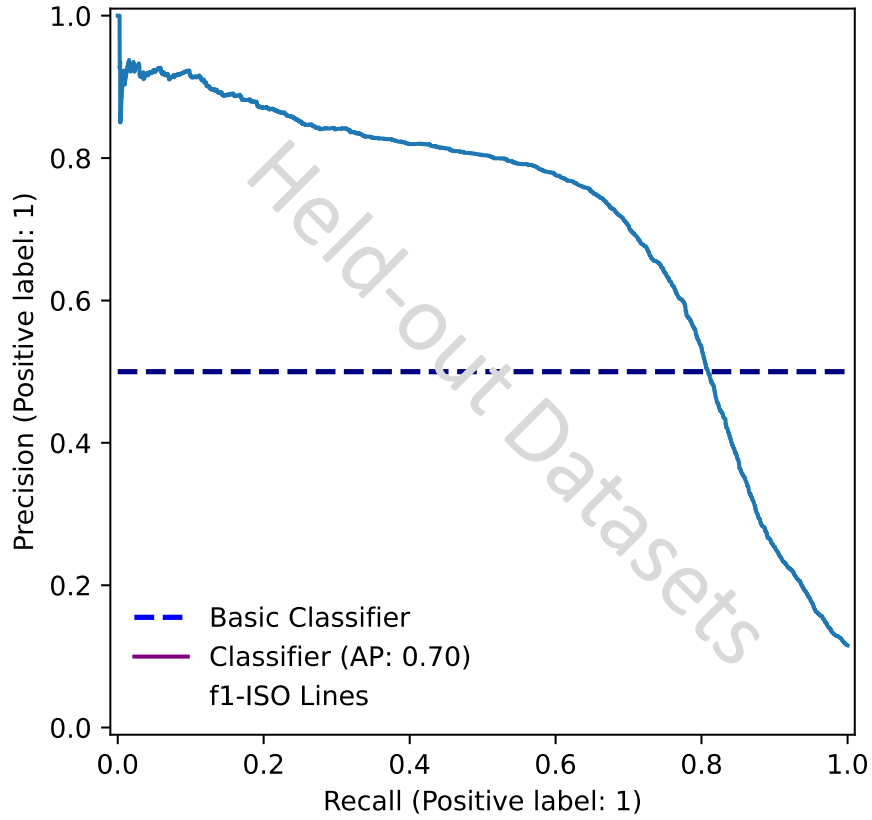
- 10,000 pages from Birds of the World
- 35,000 pages from Plants of the Worlds Online
- 5,000 species from Wikipedia

Trained on a total of more than 2,25 million samples:

- About 1,7 millions that corresponds to non-description text
- About 550,000 samples correspond to descriptive text

Model Performance





Apple blossom

Blossoms are produced in spring simultaneously with the budding of the leaves and are produced on spurs and some long shoots. The 3 to 4 cm (1 to 1+½ in) flowers are white with a pink tinge that gradually fades, five petaled, with an inflorescence consisting of a cyme with 4â6 flowers. The central flower of the inflorescence is called the "king bloom"; it opens first and can develop a larger fruit.[5][6]

The fruit is a pome that matures in late summer or autumn, and cultivars exist in a wide range of sizes. Commercial growers aim to produce an apple that is 7 to 8.5 cm (2+¾ to 3+¼ in) in diameter, due to market preference. Some consumers, especially those in Japan, prefer a larger apple, while apples below 5.5 cm (2+¼ in) are generally used for making juice and have little fresh market value. The skin of ripe apples is generally red, yellow, green, pink, or russetted, though many bi- or tri-colored cultivars may be found.[7] The skin may also be wholly or partly russeted i.e. rough and brown. The skin is covered in a protective layer of epicuticular wax.[8] The exocarp (flesh) is generally pale yellowish-white,[7] though pink, yellow or green exocarps also occur.[9]

Wild ancestors

Main article: Malus sieversii

The original wild ancestor of Malus domestica was Malus sieversii, found growing wild in the mountains of Central Asia in southern Kazakhstan, Kyrgyzstan, Tajikistan, and northwestern China.[5][10] Cultivation of the species, most likely beginning on the forested flanks of the Tian Shan mountains, progressed over a long period of time and permitted secondary introgression of genes from other species into the open-pollinated seeds. Significant exchange with Malus sylvestris, the crabapple, resulted in current populations of apples being more related to crabapples than to the more morphologically similar progenitor Malus sieversii. In strains without recent admixture the contribution of the latter predominates.[11][12][13]

Genome

Apples are diploid (though triploid cultivars are not uncommon), has 17 chromosomes and an estimated genome size of approximately 650 Mb. Several whole genome sequences have been made available, the first one in 2010 was based on the diploid cultivar 'Golden Delicious'.[14] However, this first whole genome sequence turned out to contain several errors[15] in part owing to the high degree of heterozygosity in diploid apples which, in combination with an ancient genome duplication, complicated the assembly. Recently, double- and trihaploid individuals have been sequenced, yielding whole genome sequences of higher quality.[16][17] The first whole genome assembly was estimated to contain around 57,000 genes,[14] though the more recent genome sequences support more moderate estimates between 42,000 and 44,700 protein-coding genes.[16][17] Among other things, the availability of whole genome sequences has provided evidence that the wild ancestor of the cultivated apple most likely is Malus sieversii. Re-sequencing of multiple accessions has supported this, while also suggesting extensive introgression from Malus sylvestris following domestication.[18]

Apple blossom

Blossoms are produced in spring simultaneously with the budding of the leaves and are produced on spurs and some long shoots. The 3 to 4 cm (1 to 1+½ in) flowers are white with a pink tinge that gradually fades, five petaled, with an inflorescence consisting of a cyme with 4â6 flowers. The central flower of the inflorescence is called the "king bloom"; it opens first and can develop a larger fruit.[5][6]

The fruit is a pome that matures in late summer or autumn, and cultivars exist in a wide range of sizes. Commercial growers aim to produce an apple that is 7 to 8.5 cm (2+¾ to 3+¼ in) in diameter, due to market preference. Some consumers, especially those in Japan, prefer a larger apple, while apples below 5.5 cm (2+¼ in) are generally used for making juice and have little fresh market value. The skin of ripe apples is generally red, yellow, green, pink, or russetted, though many bi- or tri-colored cultivars may be found.[7] The skin may also be wholly or partly russeted i.e. rough and brown. The skin is covered in a protective layer of epicuticular wax.[8] The exocarp (flesh) is generally pale yellowish-white,[7] though pink, yellow or green exocarps also occur.[9]

Wild ancestors

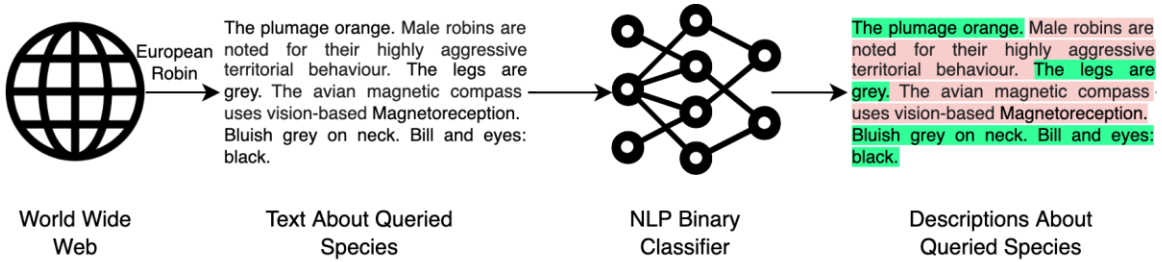
Main article: Malus sieversii

The original wild ancestor of Malus domestica was Malus sieversii, found growing wild in the mountains of Central Asia in southern Kazakhstan, Kyrgyzstan, Tajikistan, and northwestern China.[5][10] Cultivation of the species, most likely beginning on the forested flanks of the Tian Shan mountains, progressed over a long period of time and permitted secondary introgression of genes from other species into the open-pollinated seeds. Significant exchange with Malus sylvestris, the crabapple, resulted in current populations of apples being more related to crabapples than to the more morphologically similar progenitor Malus sieversii. In strains without recent admixture the contribution of the latter predominates.[11][12][13]

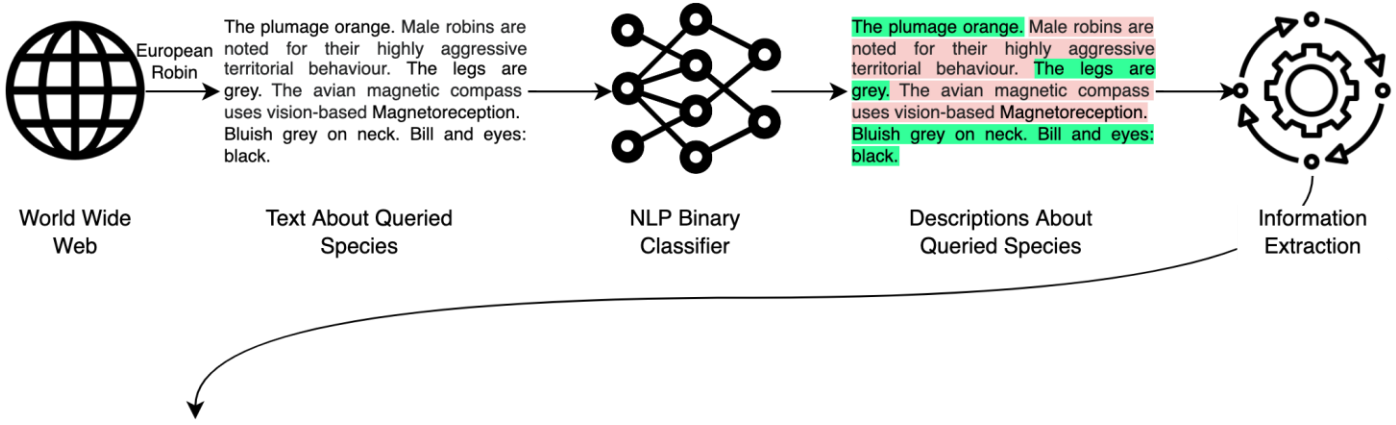
Genome

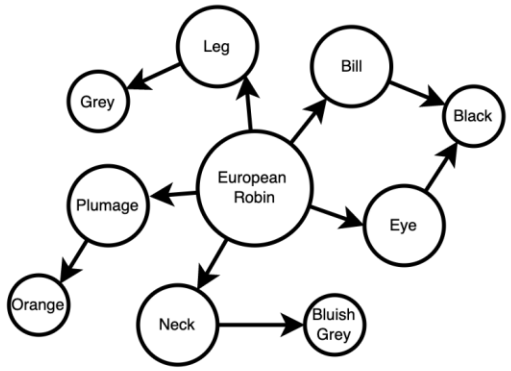
Apples are diploid (though triploid cultivars are not uncommon), has 17 chromosomes and an estimated genome size of approximately 650 Mb. Several whole genome sequences have been made available, the first one in 2010 was based on the diploid cultivar 'Golden Delicious'.[14] However, this first whole genome sequence turned out to contain several errors[15] in part owing to the high degree of heterozygosity in diploid apples which, in combination with an ancient genome duplication, complicated the assembly. Recently, double- and trihaploid individuals have been sequenced, yielding whole genome sequences of higher quality.[16][17] The first whole genome assembly was estimated to contain around 57,000 genes,[14] though the more recent genome sequences support more moderate estimates between 42,000 and 44,700 protein-coding genes.[16][17] Among other things, the availability of whole genome sequences has provided evidence that the wild ancestor of the cultivated apple most likely is Malus sieversii. Re-sequencing of multiple accessions has supported this, while also suggesting extensive introgression from Malus sylvestris following domestication.[18]

Workflow


European Robin
The plumage orange. Male robins are noted for their highly aggressive territorial behaviour. The legs are grey. The avian magnetic compass uses vision-based Magnetoreception. Bluish grey on neck. Bill and eyes: black.
The plumage orange. Male robins are noted for their highly aggressive territorial behaviour. The legs are grey. The avian magnetic compass uses vision-based Magnetoreception. Bluish grey on neck. Bill and eyes: black.
World Wide Web
Text About Queried Species
NLP Binary Classifier
Descriptions About Queried Species

Workflow





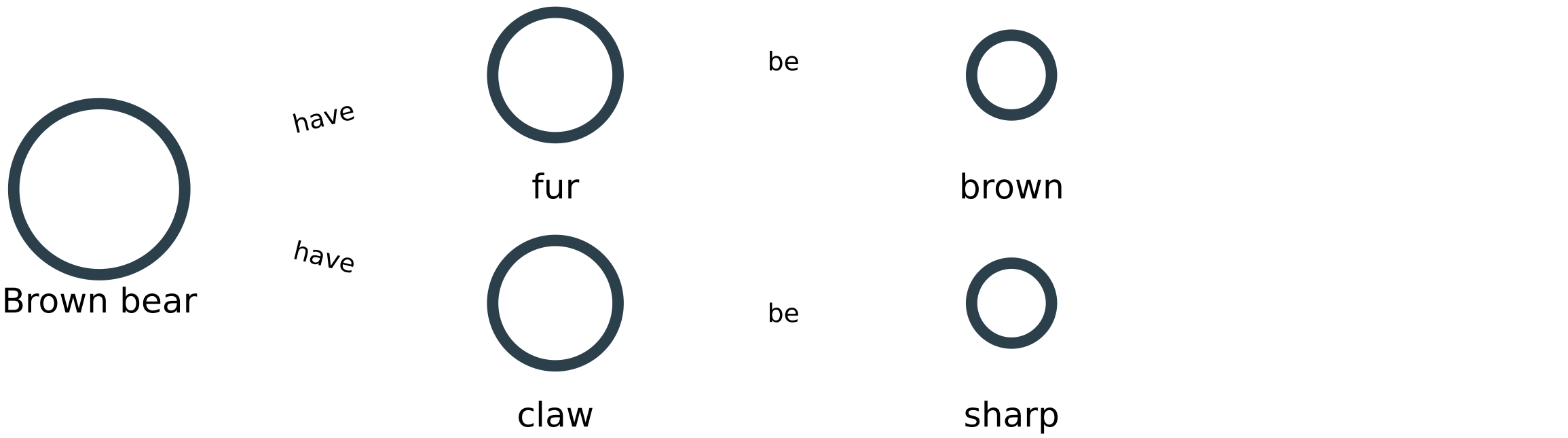
Unweighted Knowledge Graph

Workflow | Part-of-Speech Tagging

“The Brown bear has brown fur.”

“The claws are sharp.”

Workflow | Part-of-Speech Tagging



Workflow | Part-of-Speech Tagging

Does this solve potential copyright issues if the dataset is published?

Copyright protection for factual works is narrow, covering the author's original expressions, but not the facts or theories being expressed. In order to infringe, the copy must be "verbatim reproduction or very close paraphrasing".

Nonfiction literary works, such as history books, newspaper articles, and biographies, are treated as factual works with similarly narrow copyright protection. An author's unique expressions are protected, but not the facts and theories themselves. Even the selection and arrangement of facts may not be protectable.

Workflow | Part-of-Speech Tagging

73	Ceiba pentandra	86	Species has leaf. Leaf has petiolule. Petiolule is millimeter.
73	Ceiba pentandra	87	Species has leaf. Leaf has rachis. Rachis is centimeter. Centimeter is 20.
73	Ceiba pentandra	88	Species has leaf. Leaf has rachis. Rachis is centimeter. Centimeter is 5.
73	Ceiba pentandra	89	Species has leaf. Leaf has single leaf.
73	Ceiba pentandra	90	Species has leaflet. Leaflet has green. Green is dark.
73	Ceiba pentandra	91	Species has leaflet. Leaflet has petiolule. Petiolule is millimeter. Millimeter is 3.
73	Ceiba pentandra	92	Species has leaflet. Leaflet has petiolule. Petiolule is millimeter. Millimeter is 8.
73	Ceiba pentandra	93	Species has leaflet. Leaflet is 20. 20 is up.
73	Ceiba pentandra	94	Species has leaflet. Leaflet is 5.
73	Ceiba pentandra	95	Species has leaflet. Leaflet is 9.
73	Ceiba pentandra	96	Species has leaflet. Leaflet is five.
73	Ceiba pentandra	97	Species has leaflet. Leaflet is seven.
73	Ceiba pentandra	98	Species has palea. Palea has margin. Margin has nerve. Nerve has pair. Pair is 14.
73	Ceiba pentandra	99	Species has palea. Palea has margin. Margin has nerve. Nerve has pair. Pair is 5.
73	Ceiba pentandra	100	Species has palea. Palea has margin. Margin has nerve. Nerve is chartaceous.
73	Ceiba pentandra	101	Species has palea. Palea has margin. Margin has nerve. Nerve is entire.
73	Ceiba pentandra	102	Species has palea. Palea has margin. Margin has nerve. Nerve is glabrous.
73	Ceiba pentandra	103	Species has palea. Palea has margin. Margin has nerve. Nerve is pinnate. Pinnate is prominent.
73	Ceiba pentandra	104	Species has pedicel. Pedicel is long.
73	Ceiba pentandra	105	Species has pedicel. Pedicel is stout.

Workflow | Part-of-Speech Tagging

How do these compare to manually curated databases?

PalmTraits 1.0, a species-level functional trait database of palms worldwide

W. Daniel Kissling, Henrik Balslev, William J. Baker, John Dransfield, Bastian Göldel, Jun Ying Lim, Renske E. Onstein & Jens-Christian Svenning

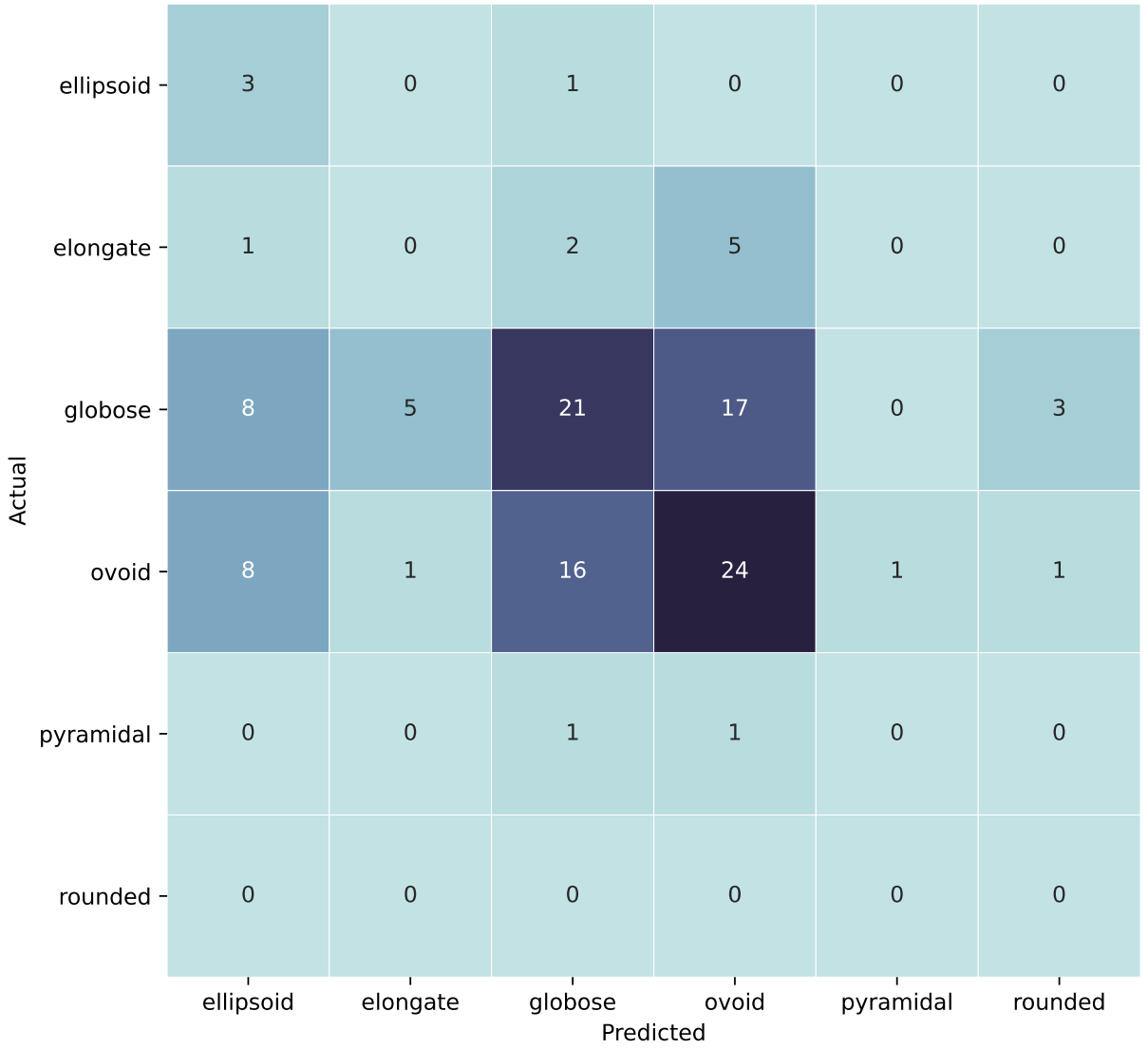
Scientific Data **6**, Article number: 178 (2019) | Cite this article

SpecName	Acanthophoenix crinita
Erect	1
StemSolitary	1
StemArmed	1
MaxStemHeight_m	10
MaxStemDia_cm	20
MaxLeafNumber	15
AverageFruitWidth_cm	0.5
FruitSizeCategorical	small
FruitShape	NaN
FruitColorDescription	black
MainFruitColors	black

Palm Trait Dataset Binary Class Variables

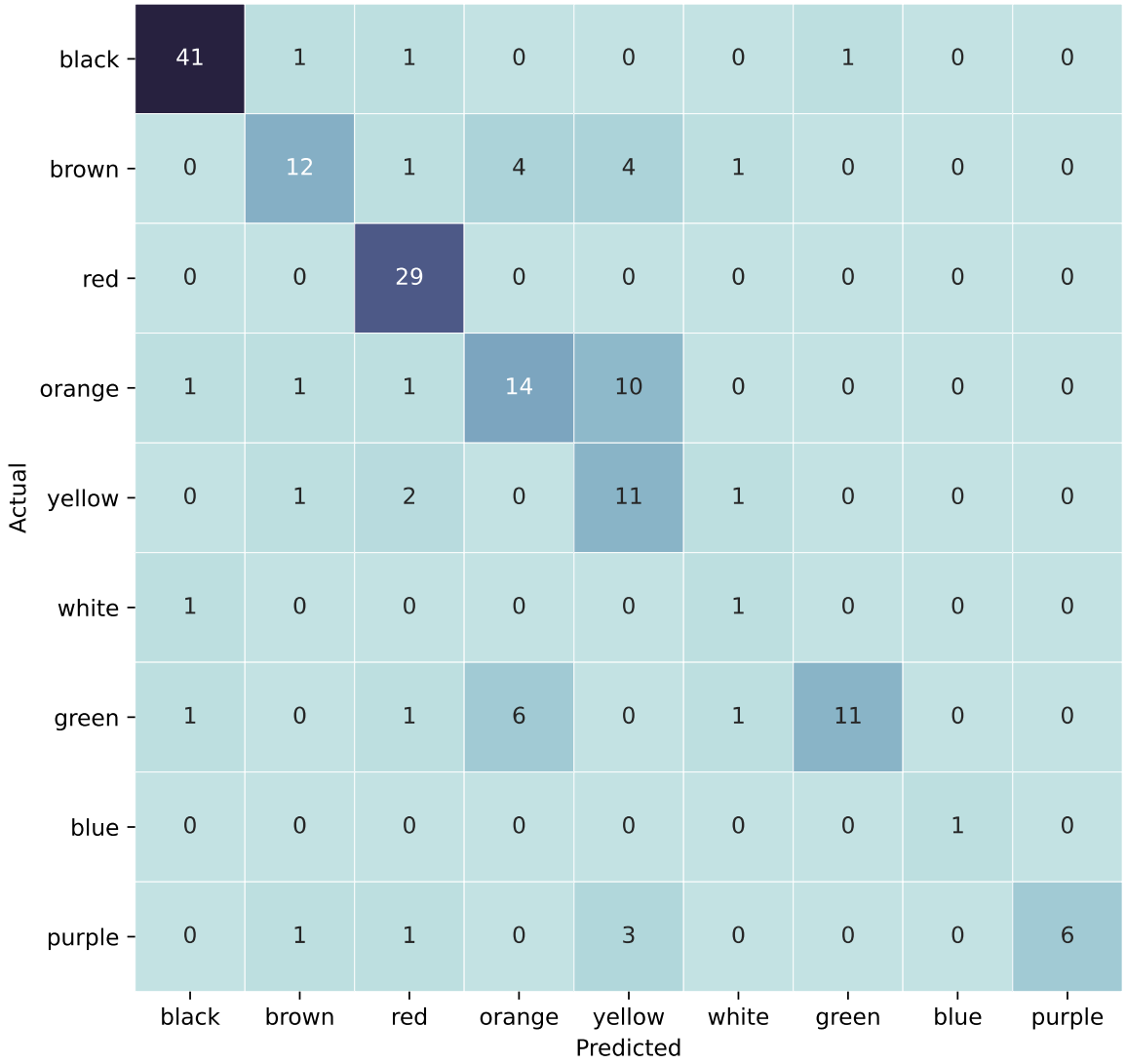
	Precision	Recall	F1 Score
Climbing	0.952	0.976	0.964
Acaulescent	0.905	0.948	0.926
StemErect	0.924	0.452	0.560
StemSolitary	0.833	0.681	0.714
StemArmed	0.838	0.911	0.873

Multi Class Variables Shapes

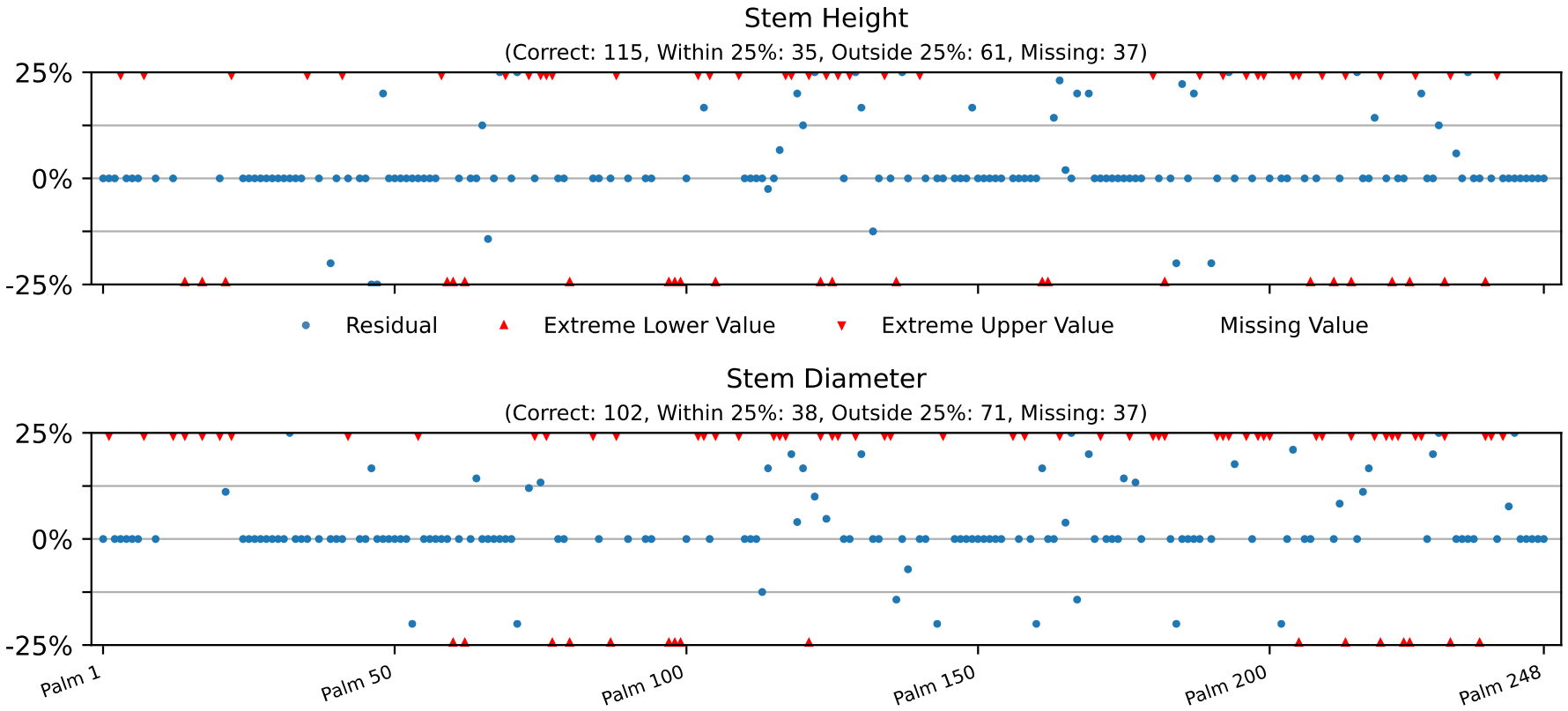


	Missing	Precision	Recall	F1 Score
Shapes	129	0.450	0.193	0.267
Colours	77	0.770	0.508	0.595

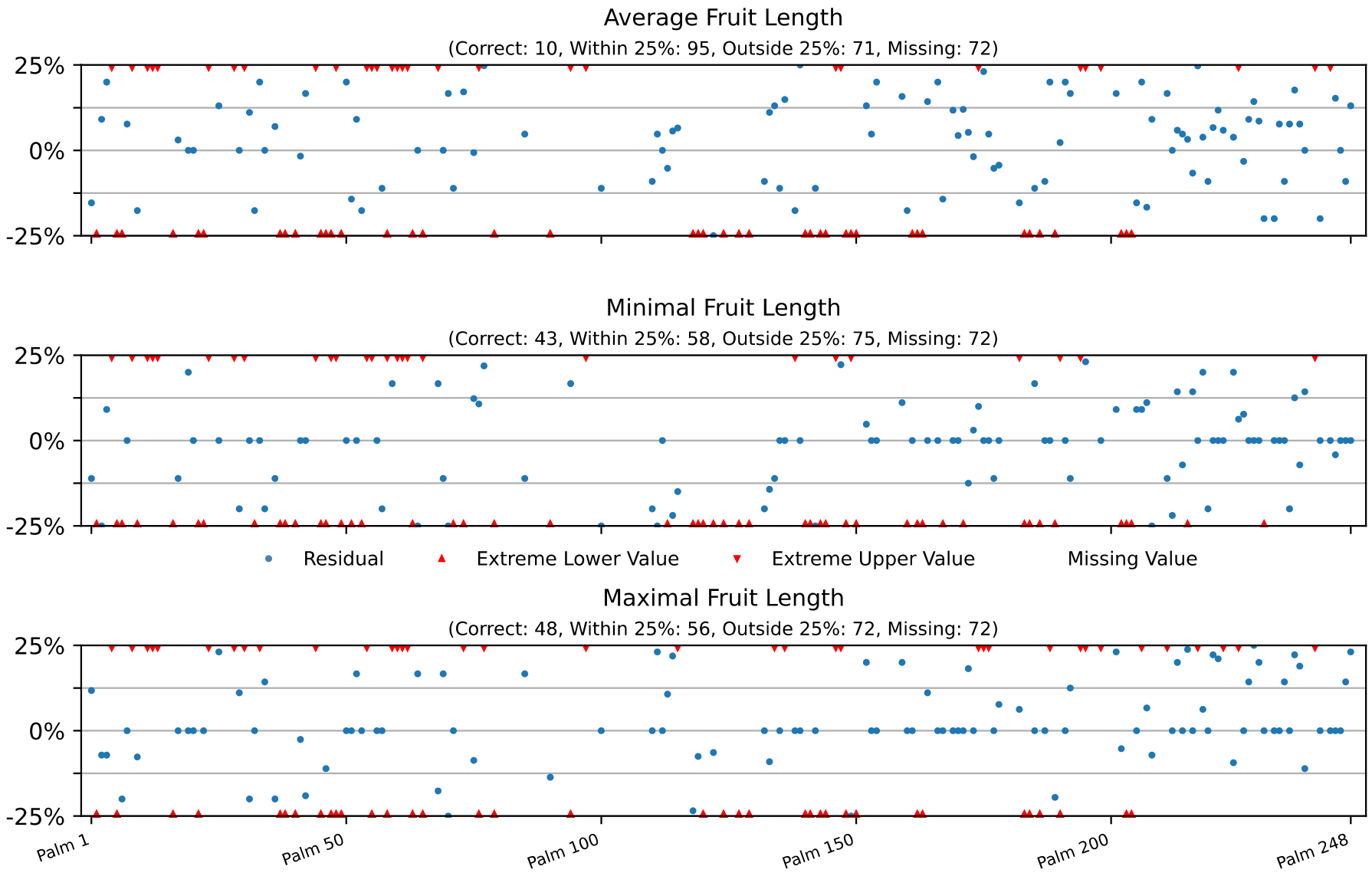
Multi Class Variables Colors

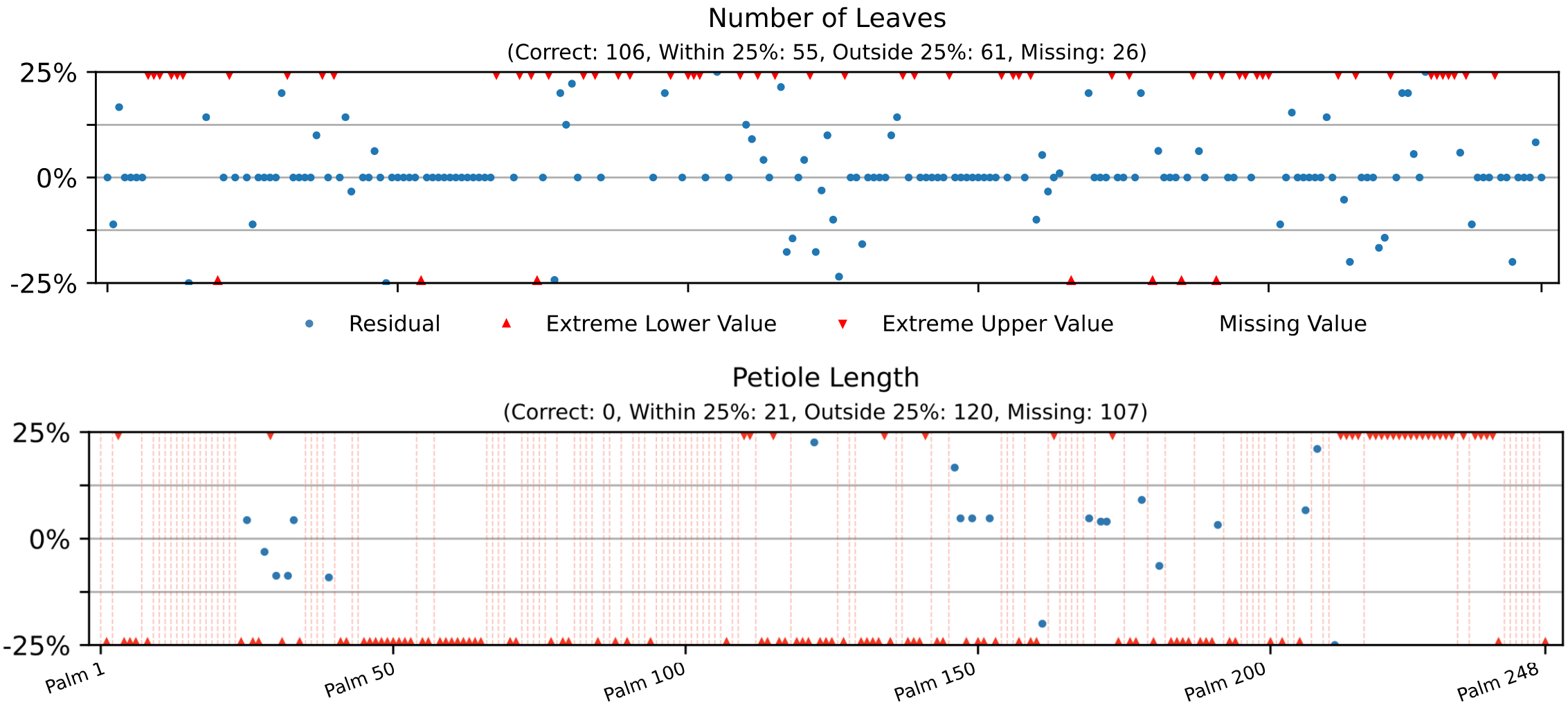


	Missing	Precision	Recall	F1 Score
Shapes	129	0.450	0.193	0.267
Colours	77	0.770	0.508	0.595



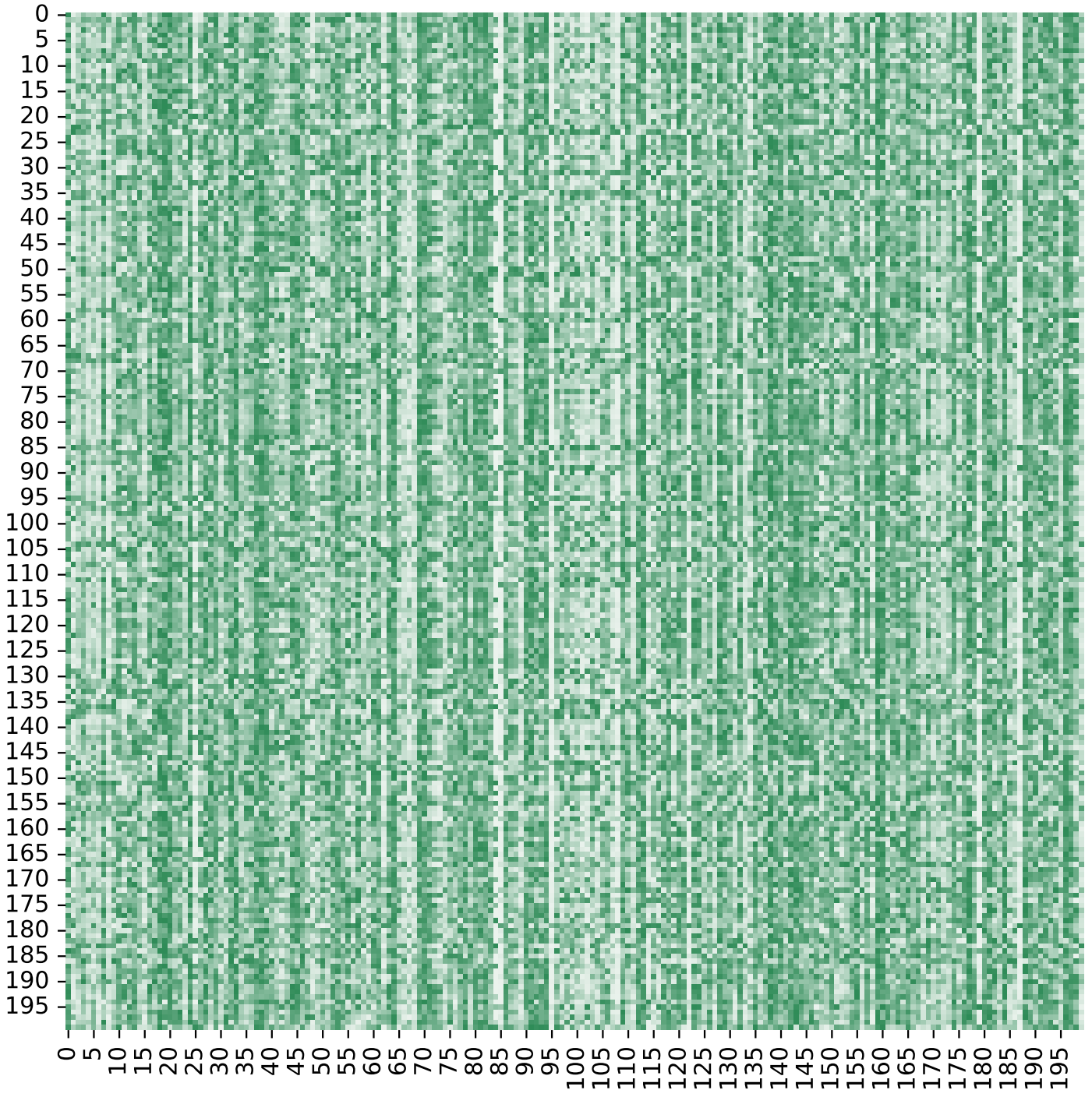
Palm Trait Dataset
Numerical Values


Average Fruit Length
(Correct: 10, Within 25%: 95, Outside 25%: 71, Missing: 72)
25%
0%
-25%
Minimal Fruit Length
(Correct: 43, Within 25%: 58, Outside 25%: 75, Missing: 72)
25%
0%
-25%
Residual
Extreme Lower Value
Extreme Upper Value
Missing Value
Maximal Fruit Length
(Correct: 48, Within 25%: 56, Outside 25%: 72, Missing: 72)
25%
0%
-25%
Palm 1
Palm 50
Palm 100
Palm 150
Palm 200
Palm 248


Number of Leaves
(Correct: 106, Within 25%: 55, Outside 25%: 61, Missing: 26)
25%
0%
-25%
Residual
Extreme Lower Value
Extreme Upper Value
Missing Value
Petiole Length
(Correct: 0, Within 25%: 21, Outside 25%: 120, Missing: 107)
25%
0%
-25%
Palm 1
Palm 50
Palm 100
Palm 150
Palm 200
Palm 248

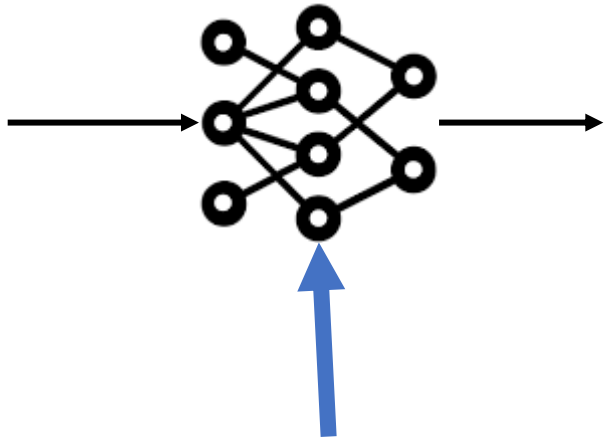
CUB-200 Dataset Ranking Comparison

Average Ranking 40.15! (out of 200 ranked birds)



Next steps | Path A: tabular data

Species has fruit. Fruit has green.
Species has fruit. Fruit has layer of fruit pulp. Layer of fruit pulp is thick.
Species has fruit. Fruit has maturation. Maturation has fruit maturation.
Species has fruit. Fruit has pericarp. Pericarp is brown. Brown is reddish.
Species has fruit. Fruit has pericarp. Pericarp is brown. Brown is yellowish.
Species has fruit. Fruit has pericarp. Pericarp is fleshy,.sugary.
Species has fruit. Fruit has pulp. Pulp has fruit pulp.
Species has fruit. Fruit has skin. Skin is smooth.
Species has fruit. Fruit has texture. Texture is firm.
Species has fruit. Fruit has tree. Tree has fruit tree.
Species has fruit. Fruit is bedraya.
Species has fruit. Fruit is black.
Species has fruit. Fruit is bloom. Bloom is like. Like is frost.



SpecName	Acanthophoenix crinita
Erect	1
StemSolitary	1
StemArmed	1
MaxStemHeight_m	10
MaxStemDia_cm	20
MaxLeafNumber	15
AverageFruitWidth_cm	0.5
FruitSizeCategorical	small
FruitShape	NaN
FruitColorDescription	black
MainFruitColors	black

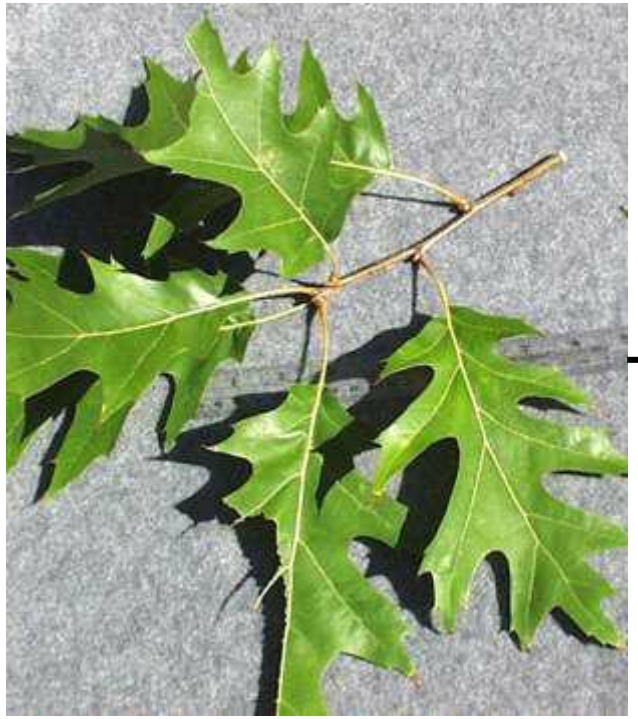
Supervised with existing curated databases

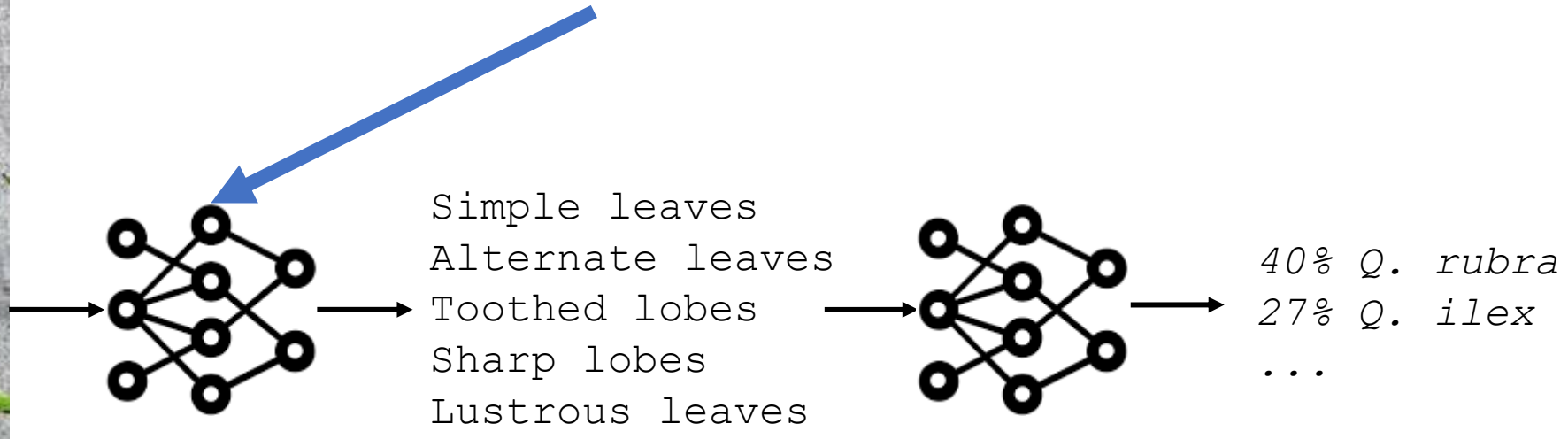
Next steps | Path B: refine the text

- Clean irrelevant information
- Homogenize traits across species
- Focus on observable traits

Next steps | Visual-text Hybrid model

How to train this? We only have a class label per image!





Next steps | Visual-text Hybrid model

Simple leaves
Alternate leaves
Toothed lobes
Sharp lobes
Lustrous leaves

Scaly bark
Dark reddish gray
Smooth when young

Standard captioning task

Next steps | Visual-text Hybrid model

Q. rubra

Q. rubra

Simple leaves
Alternate leaves
Toothed lobes
Sharp lobes
Lustrous leaves
Scaly bark
Dark reddish gray
Smooth when young
Etc.
Etc.
...

Our supervision is much weaker!

Lots of work ahead!

Thank you!