



Difesa sostenibile della vite

Intelligenza Artificiale in campo: modelli predittivi e strumenti di analisi in Enophit

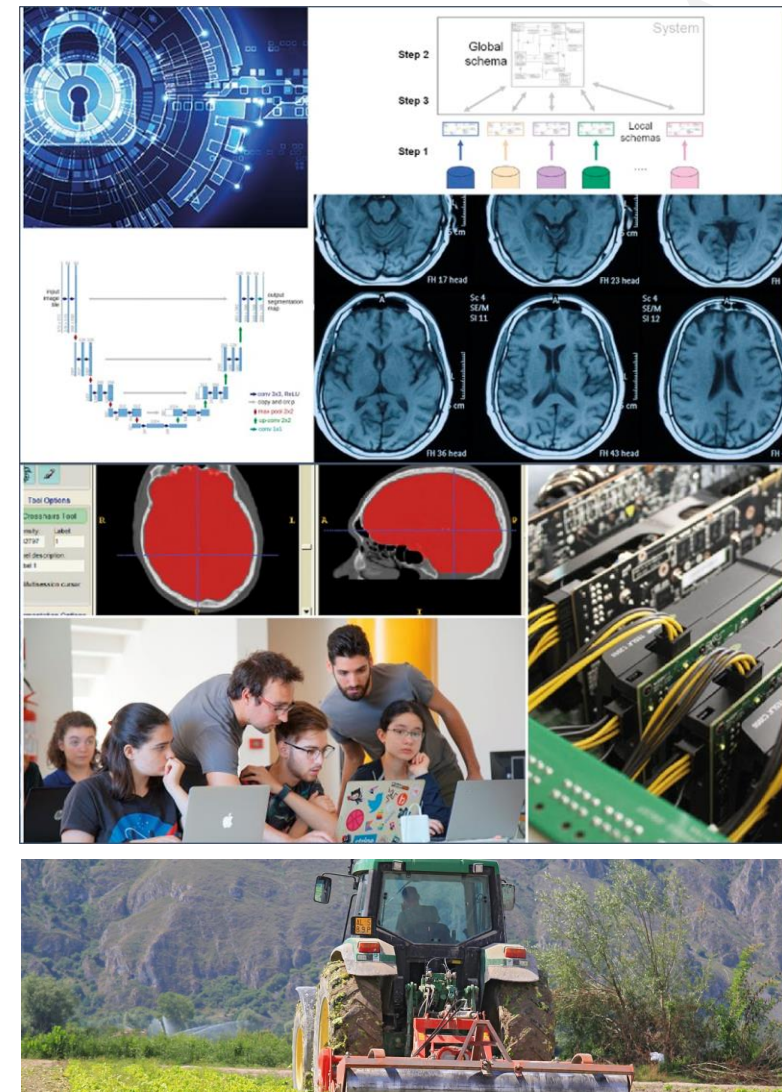
Cesare Furlanello

FBK/MPBA



Intelligenza Artificiale (IA): la nuova sfida tecnologica che sviluppa macchine con l'abilità di apprendere da sole senza essere state esplicitamente programmate

L'IA rende possibili applicazioni come la guida autonoma, il riconoscimento facciale e la traduzione automatica; tra le sfide più ambiziose ci sono ora **miglioramento della salute e agricoltura digitale**



- **LE MACCHINE APPRENDONO?**

l'apprendimento sembra un aspetto caratteristico dell'ingegno umano, ma è un processo generale, basato sulla assimilazione di esempi ed esperienze e la sintesi di informazioni

Per imparare a identificare oggetti e acquisire attività base, che combinate assieme permettono di svolgere compiti anche molto complessi, è fondamentale avere a disposizione esempi e tempo.

Un processo continuo, non solo la esecuzione di regole



Impariamo dagli esempi, spesso tramite un insegnante, modificando azioni e decisioni dai nostri errori

- IMPARARE
DA ESEMPI



Per compiti complessi (e critici !) come la guida autonoma, i sistemi di IA apprendono a riconoscere pedoni, veicoli, segnali, elementi delle strade e a anticipare azioni future. Ogni componente è addestrata per prova ed errore da esempi: migliaia di scene, in cui gli elementi sono stati etichettati. Dati nuovi, mai visti in addestramento, sono usati come test virtuale per valutare il sistema.

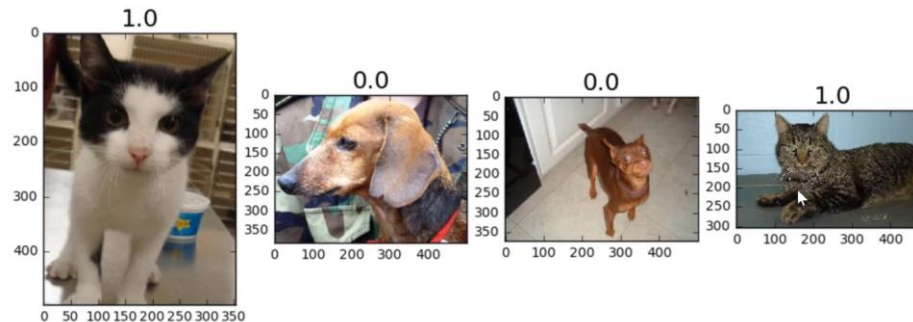
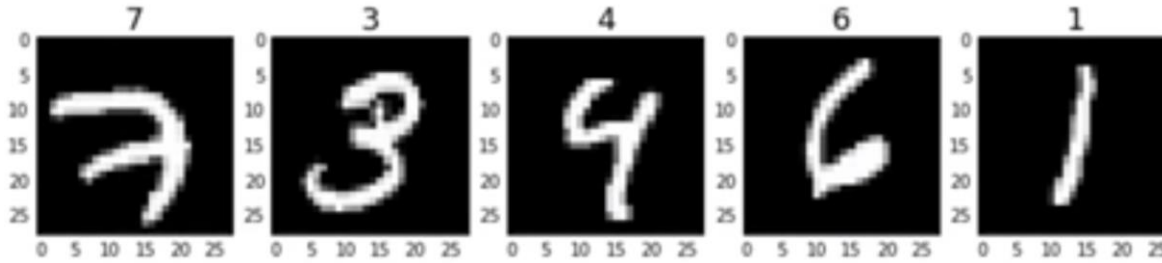
In 30''

La macchina impara a costruire la associazione tra un segnale e la sua etichetta, senza essere esplicitamente programmata, ma principalmente dagli esempi

Oggetto
(es: immagine o sequenza di dati)



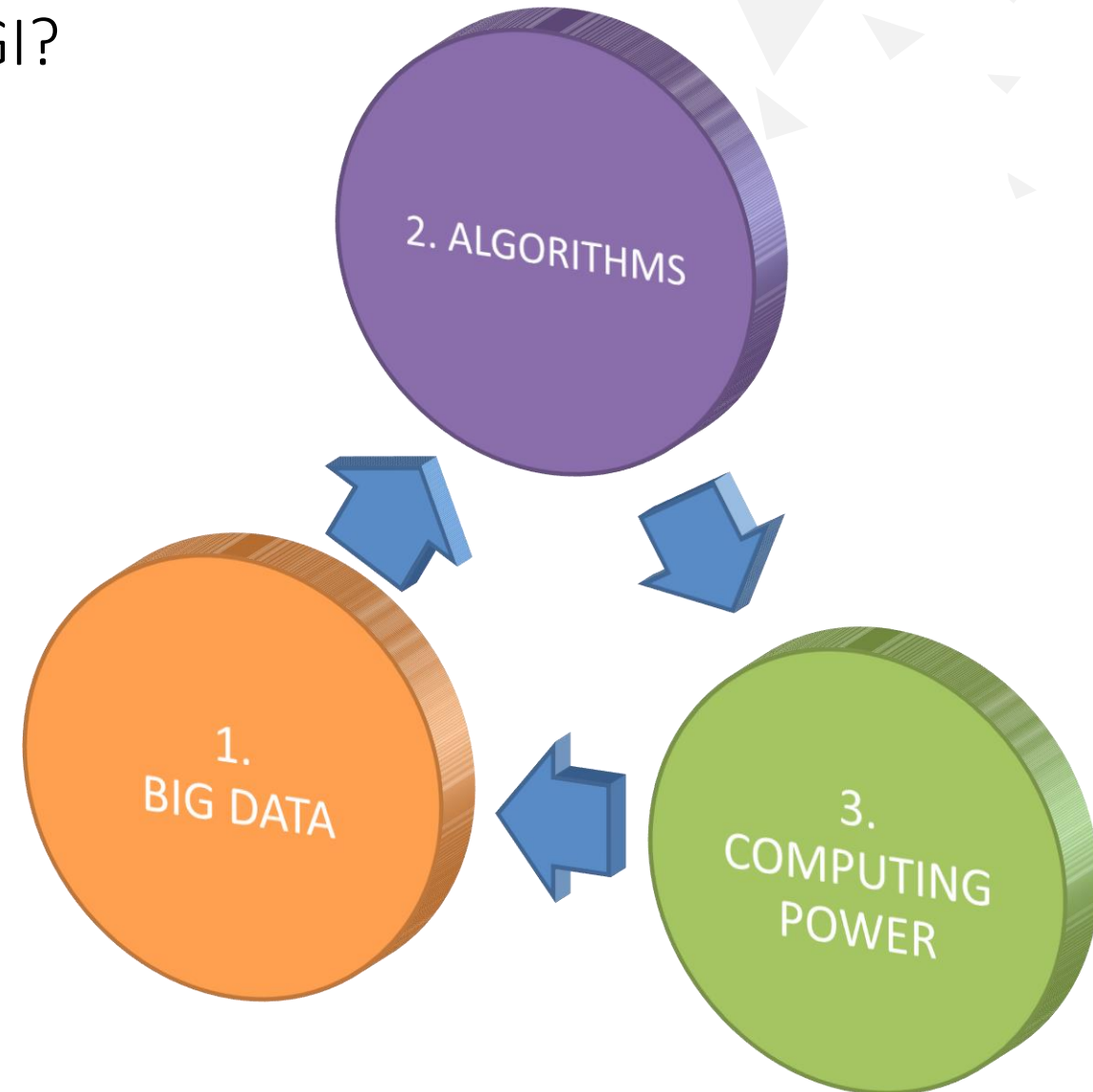
Etichetta
(classe o valore)



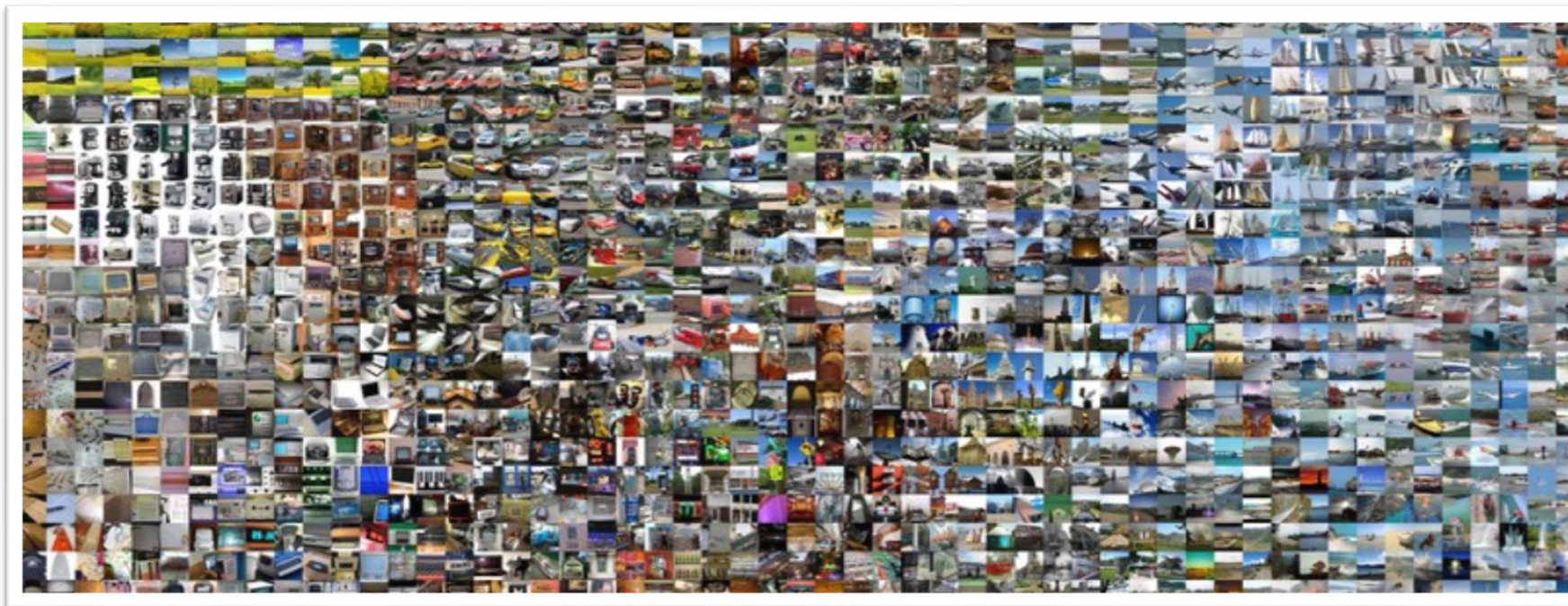
INTELLIGENZA ARTIFICIALE: PERCHE' OGGI?

Tre fattori principali si sono combinati per questo salto in avanti

1. Big Data per addestrare i modelli
2. Nuovi algoritmi neurali
3. Soluzioni di calcolo dedicate (le GPU, anche «in campo»)



“BIG DATA” per la IA: IMAGENET



- Total number of non-empty synsets: 21841
- Total number of images: 14,197,122

IMAGENET

I Big Data sono una delle tre basi della rivoluzione della IA negli ultimi 10 anni

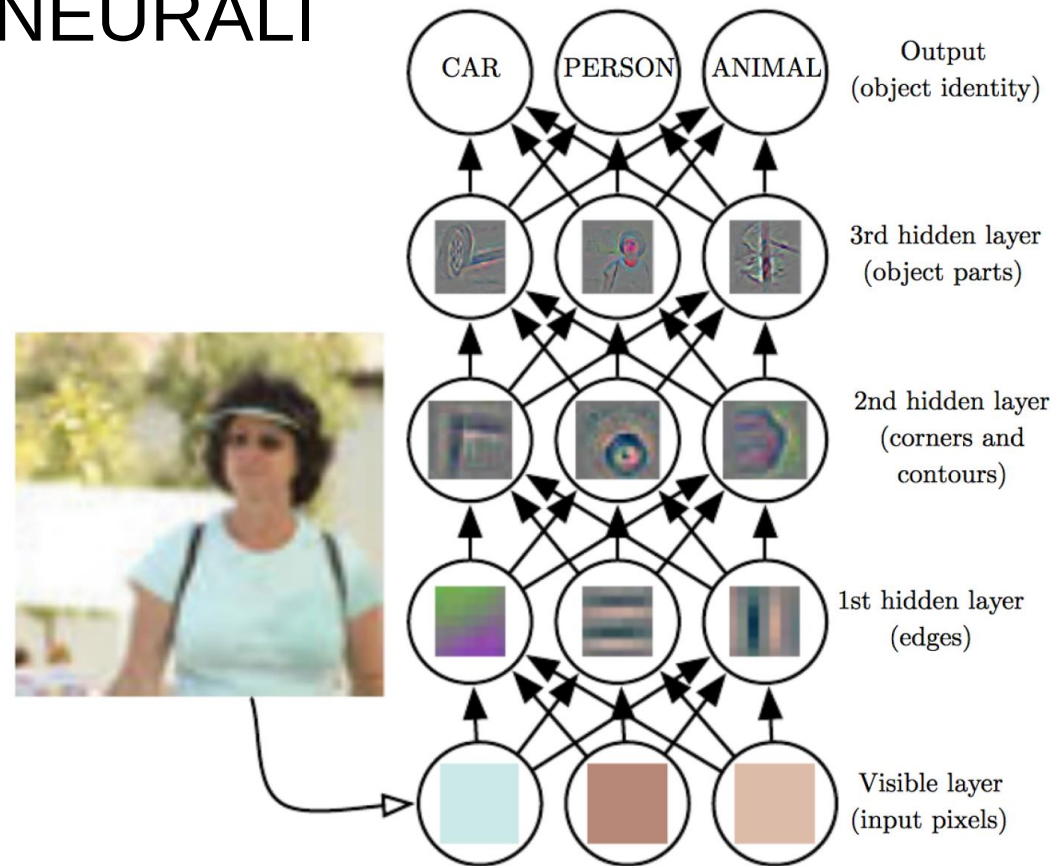
- **IMAGENET raccoglie 14 Milioni di immagini** etichettate che possono essere usate per addestrare sistemi di riconoscimento, anche per più di 20 000 classi
- Tutte le immagini «pubbliche» di Instagram, Facebook, e in pratica tutti i nostri dati in cloud, sono correntemente usate per addestrare i sistemi, senza etichette.

ALGORITMI PER APPRENDERE: RETI NEURALI

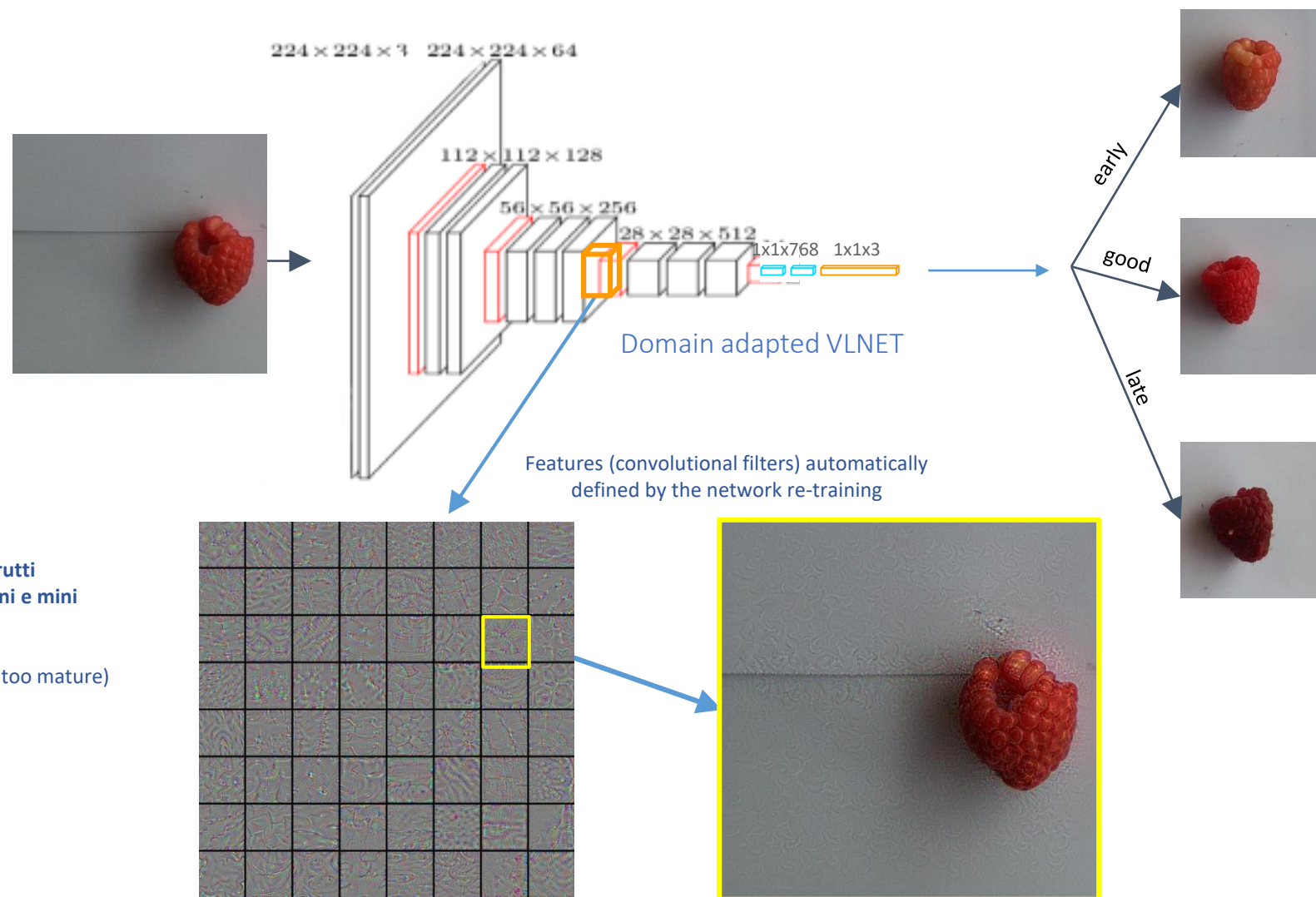
Semplificando, gli algoritmi del machine learning cercano strutture e regole funzionali tra i dati osservati per generalizzare la conoscenza

Per avvicinarsi alle capacità di apprendimento della mente umana:

- i **modelli a reti neurali addestrati con le tecniche di Deep Learning** ottengono una funzione complessa imparando a riconoscere dettagli e a combinarli assieme.
- Esempi di applicazioni: classificare, produrre descrizioni testuali di immagini, tradurre testi.



Deep Learning per Controllo Maturazione



OBIETTIVO

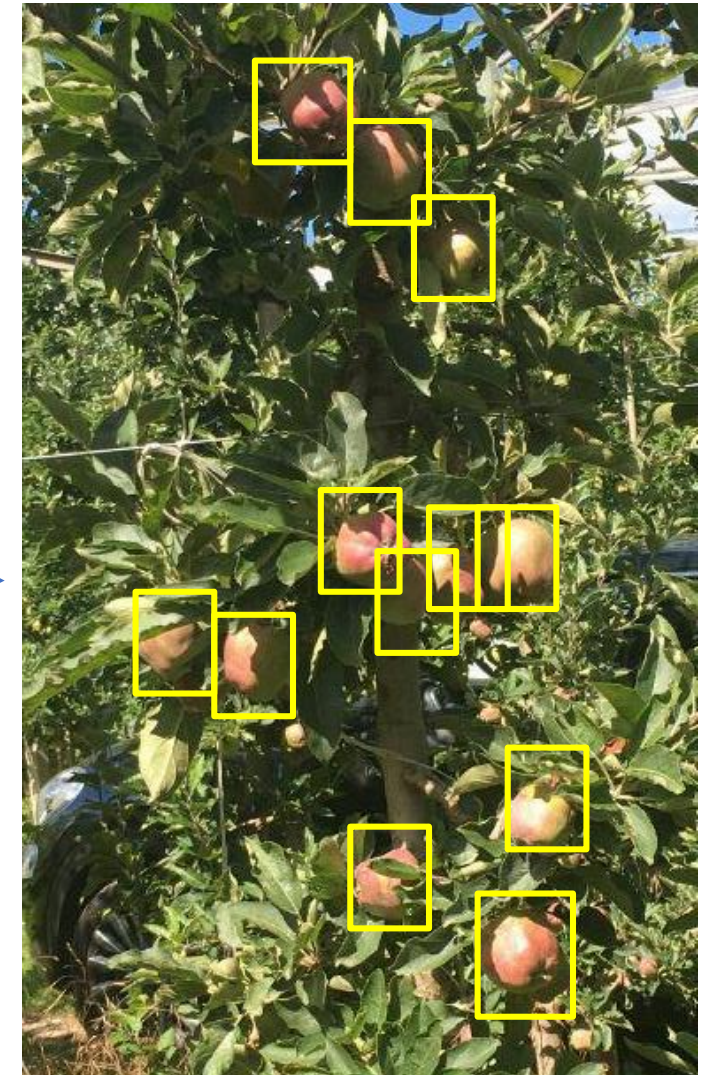
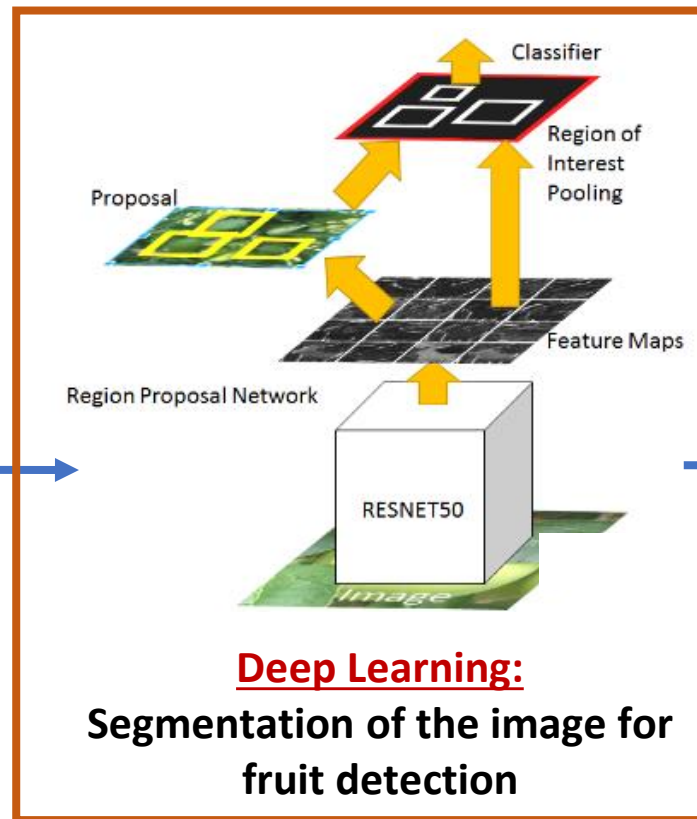
- **STIMA DELLA MATURAZIONE** di piccoli frutti utilizzando una combinazione di immagini e mini spettrometri
- **TRE CLASSI** (Early, Good for harvest, Late too mature)
- Quality Control
- Portabile, Non distruttiva
- Low-cost
- Smartphone e mobile first con soluzione cloud

Domain adaptation: una rete neurale addestrata su esempi di IMAGENET, che specializziamo con migliaia di nuove immagini addestrando a 3 classi di maturazione

AI for Fruit detection and damage assessment

Database Casez WV2017: segmented 34.000

Image acquisition by smartphone or automatic systems



Yield estimation: 2018

Annotated berries 2017/2018

Teroldego 18.000

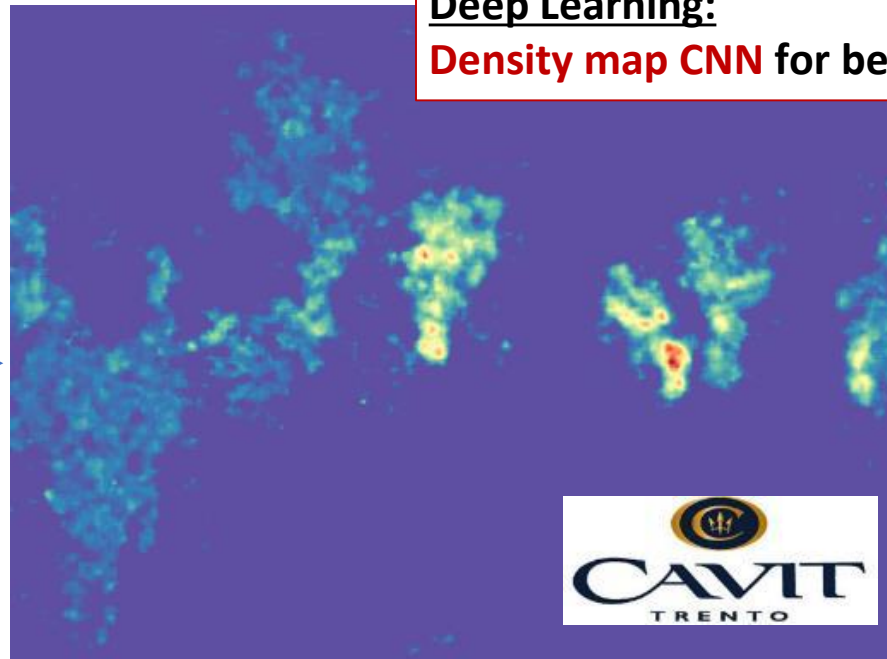
Preliminary results on Teroldego variety:

Mean average error ~10% on 8000 berries

Training images dataset:

Teroldego (2017)	67	Pinot Grigio	40
Teroldego (2018)	20	Pinot Nero	24
Lagrein	20	Sauvignon	39
Marzemino	24	Traminer	42
Merlot	20		

Image acquisition
using smartphone or
other automatic
systems



Deep Learning:

Density map CNN for berries counting



Prediction: # berries

Database
variety data

GIS / Catasto
(Cartografia)

Yield estimation

AI E DEEP LEARNING: FENOLOGIA

Sequenza
(Temp medie)

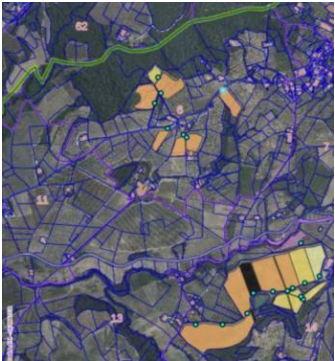
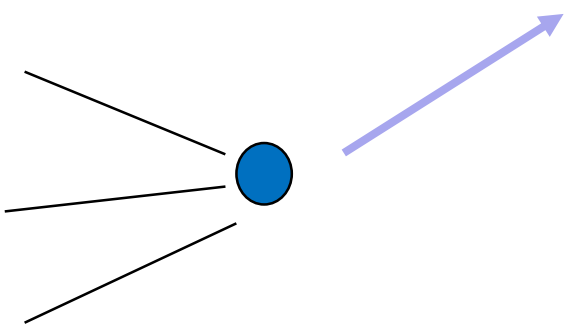


Stadio
Fenologico

Etichette = fasi fondamentali,
oppure tutte le fasi BBCH

Modelled process	Equations
Dormancy induction The rate of dormancy induction is negatively related to both temperature and photoperiod through sigmoidal relationships and bound between 0 and 1. Dormancy induction is thus completed at t_d when the state of dormancy induction (at time t^*), $DS(t) \geq Dcrit$ (critical state of dormancy induction). Dormancy induction state (DS) is defined by the daily accumulation of rates of dormancy triggered by temperature (DR_T) and photoperiod (DR_P) induction.	$DS(t) = \sum_{i=0}^t DR_T(i) \times DR_P(i)$ $DR_T(t) = \frac{1}{1 + e^{a(T_t - T_{opt})}}$ $DR_P(t) = \frac{1}{1 + e^{b(P_t - P_{opt})}}$
Chilling The rate of chilling is described by a modified version of the chilling function of the Unified Model (Chaine 2000). The state of chilling (CS) was defined by the daily accumulation of rates of chilling (CR), driven by temperature.	$CS(t) = \sum_{i=0}^t \frac{1}{1 + e^{c(T_i - T_{crit})}}$
Forcing The state of forcing (FS) is defined by the daily accumulation of the rate of forcing (FR) defined by a positive and sigmoidal relationship with temperature. $FR(t)$ is in turn dependent on critical photoperiod (DL_{crit}), which depends on the amount of accumulated chilling units at time t [$CS(t)$]. Crit defines the position of the inflection point of DL_{crit} ; hence, it can be considered as an indicator of the critical amount of exposure to chilling necessary for dormancy release (critical chilling state). Budburst is simulated on the date when $FS \geq Fcrit$ (critical forcing state).	$FS(t) = \sum_{i=0}^t \frac{1}{1 + e^{d(T_i - T_{opt})}}$ $T_{crit}(t) = \frac{60^2}{1 + e^{0.01(DL_{crit} - DL_{opt})}}$ $DL_{crit}(t) = \frac{24^2}{1 + e^{0.002(300 - GSU)}}$

D
C
F



Caffarra et al 2014

Problema: una volta definita la funzione da principi di fisiologia vegetale, vanno calibrati i parametri in funzione di differenze ambientali, di vitigno ecc
da dati molto sparsi

AI E DEEP LEARNING: FENOLOGIA

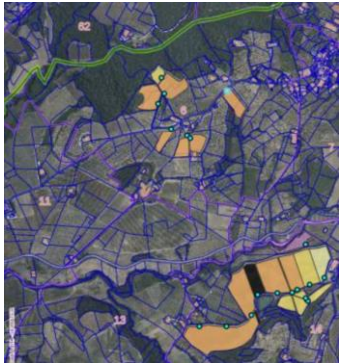
Sequenza
Temp

BBCH

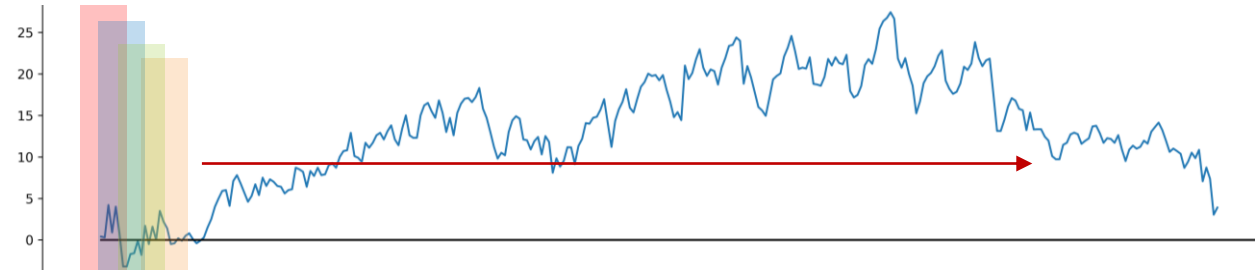
Serie di temperature
a risoluzione 2km

PER TUTTE LE PARTICELLE
COLTIVATE A VITE

Particelle georiferite

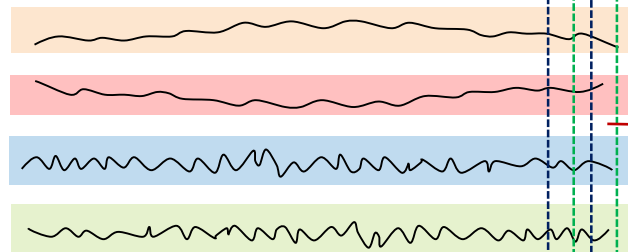


Temperatura media giornaliera

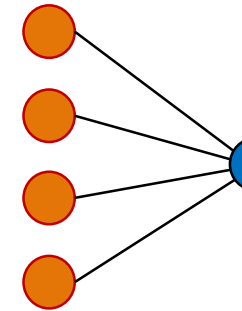


4 filtri convoluzionali scorrono
sulla serie delle temperature

2 filtri convoluzionali scorrono
sull'output dei filtri precedenti



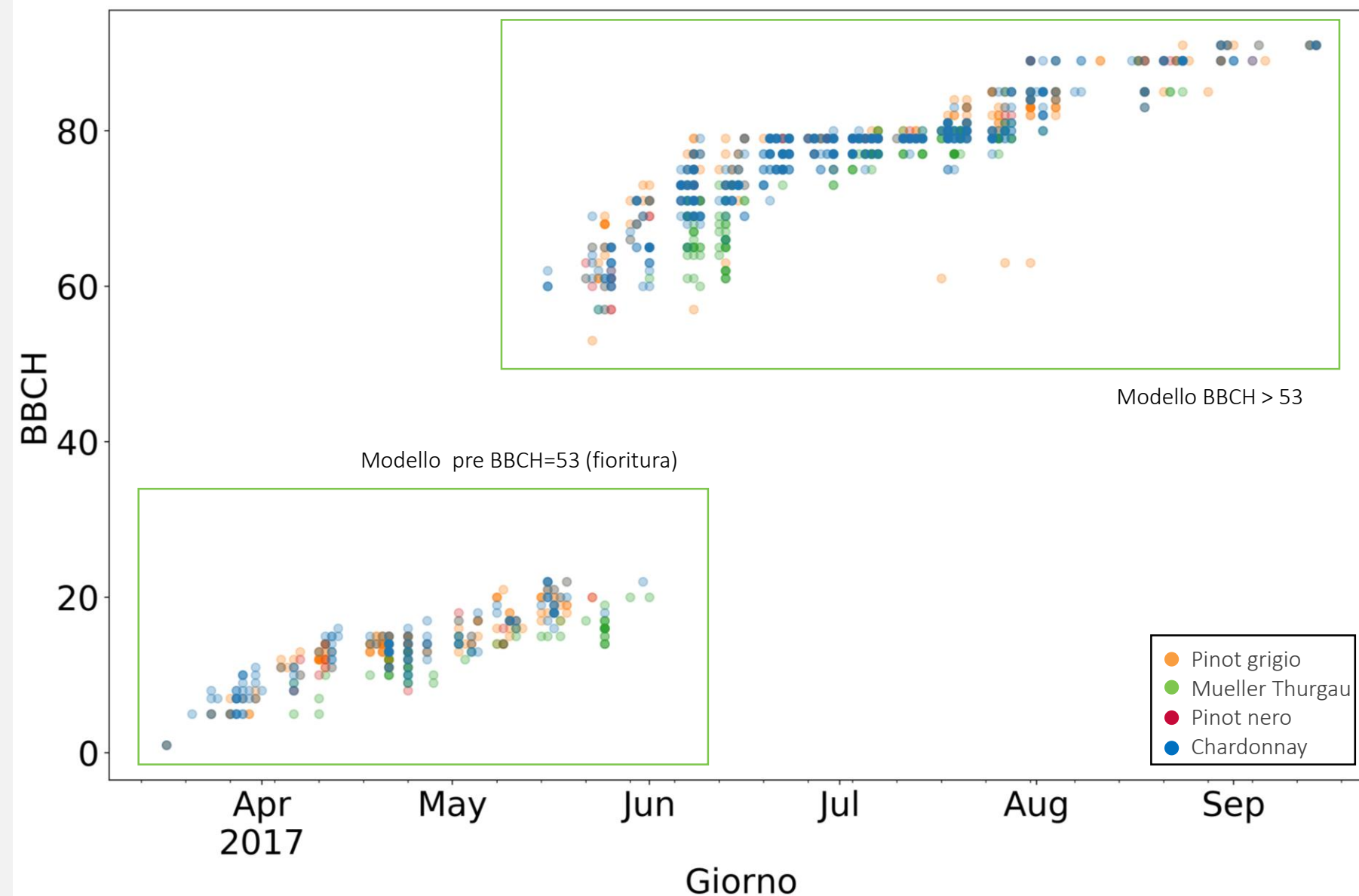
$n_chans=[4, 2]$, $ksizes=[5, 3]$, $strides=[3, 2]$



Sezione output
rete fully connected

Predizione
BBCH

DATI TRAINING PER IL MODELLO DL ($R^2=0.79$)



Rilevazioni BBCH
annotate tramite quaderni di
campagna (MPA Solutions)

Rilevazioni 2011-2016

Pinot grigio: 337
Mueller Thurgau: 267
Pinot nero 204
CHARDONNAY 1237

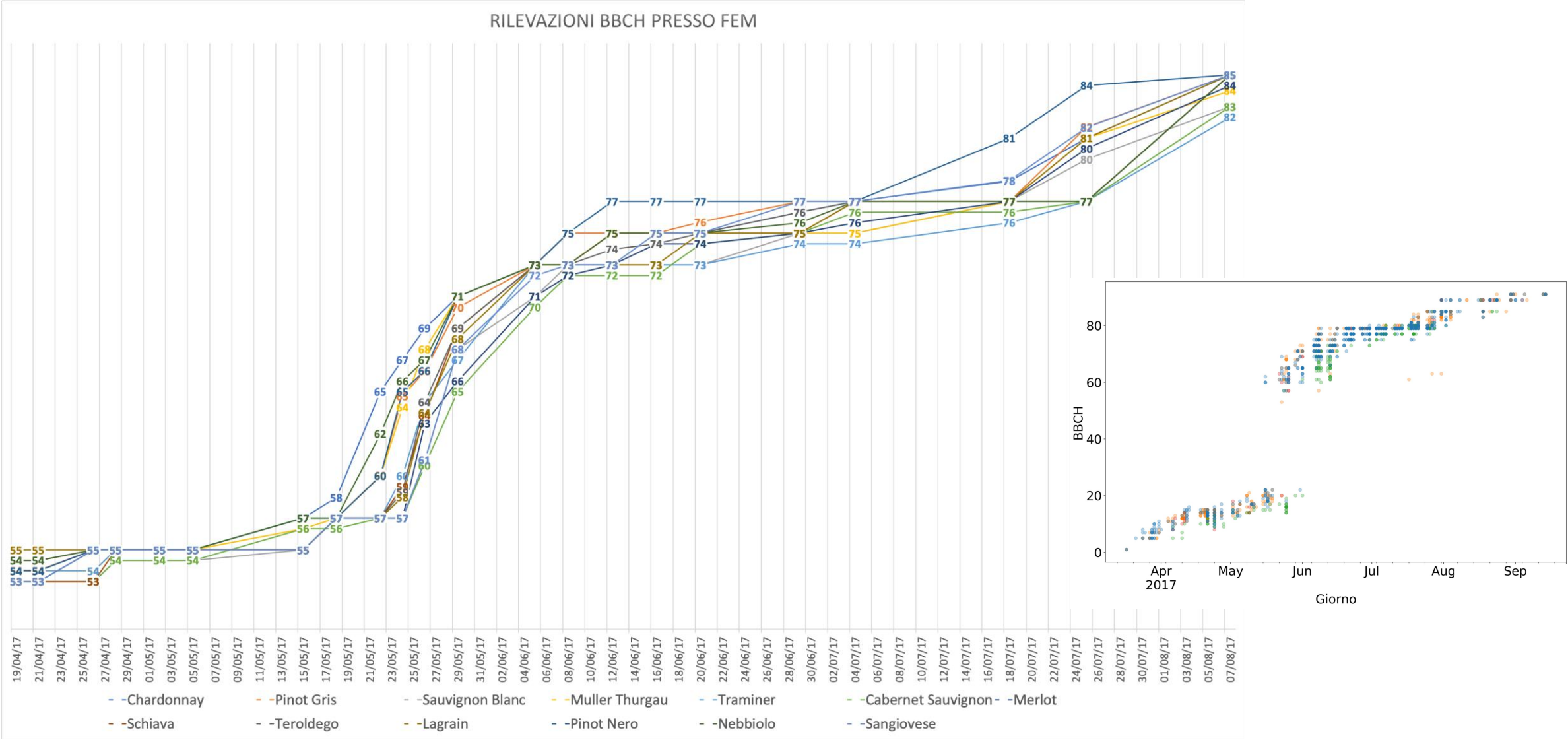
Training dataset

Rilevazioni 2017

Pinot grigio 369
Mueller Thurgau 184
Pinot nero 57
CHARDONNAY 515

Test dataset

UNA RILEVAZIONE IDEALE (FEM)



IL MODELLO DI DECISIONE (DSS)

SIMULAZIONE DSS 2017: applichiamo le “raccomandazioni”

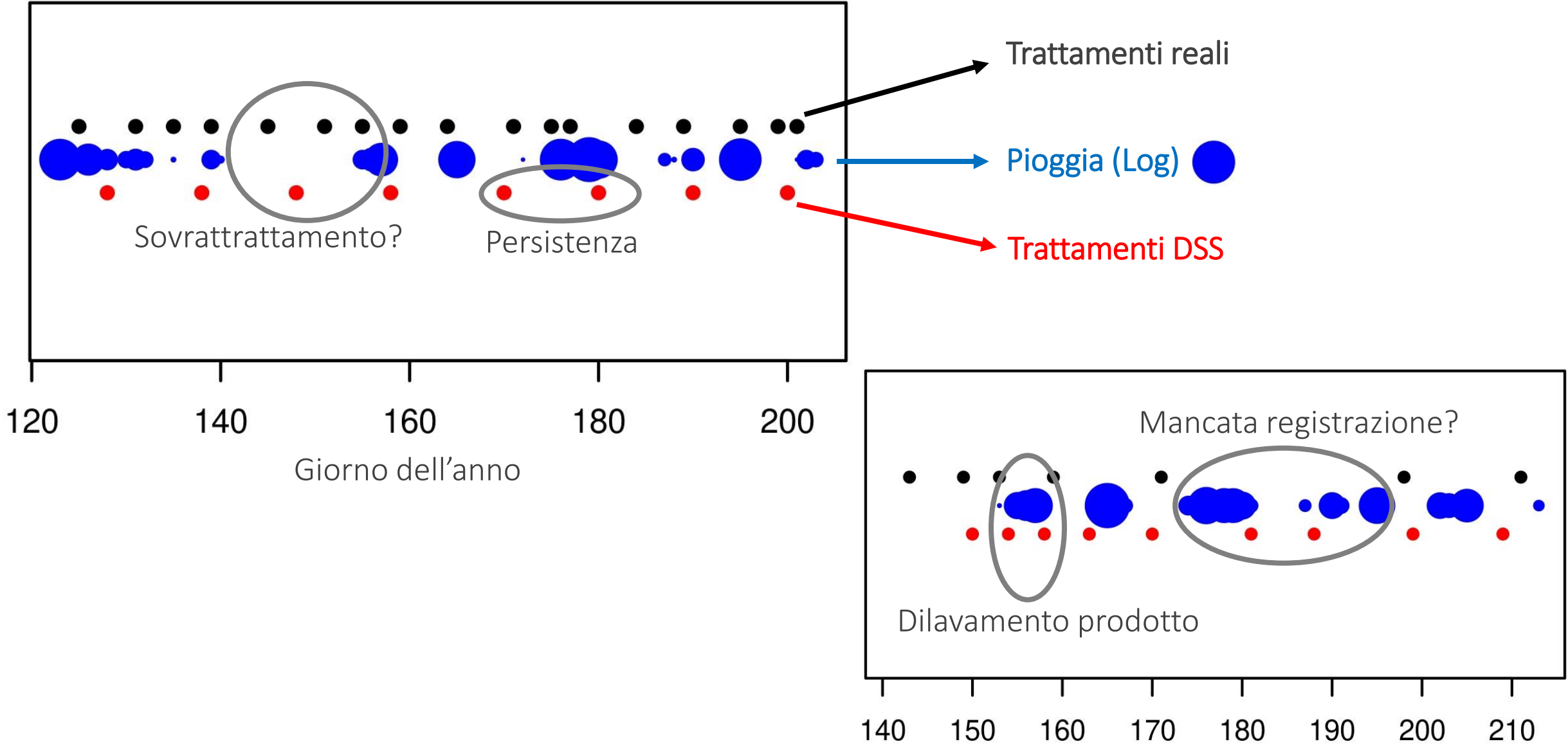
- ~4500 appezzamenti georiferiti ← dati Temp da GIS
~324K trattamenti singoli, 30 avversità
- Periodo*: 1 Gennaio a 30 Agosto
- Simulazione PERONOSPORA
- **FENOLOGIA DL + REGOLE DSS**
- 4423 appezzamenti con dati consistenti,
57476 trattamenti reali (m=12.99)
- Inizializzazione magazzini (141 prodotti utilizzati)
- Regola DSS: accettazione immediata consiglio DSS

* finestra operativa individualizzata
su appezzamento

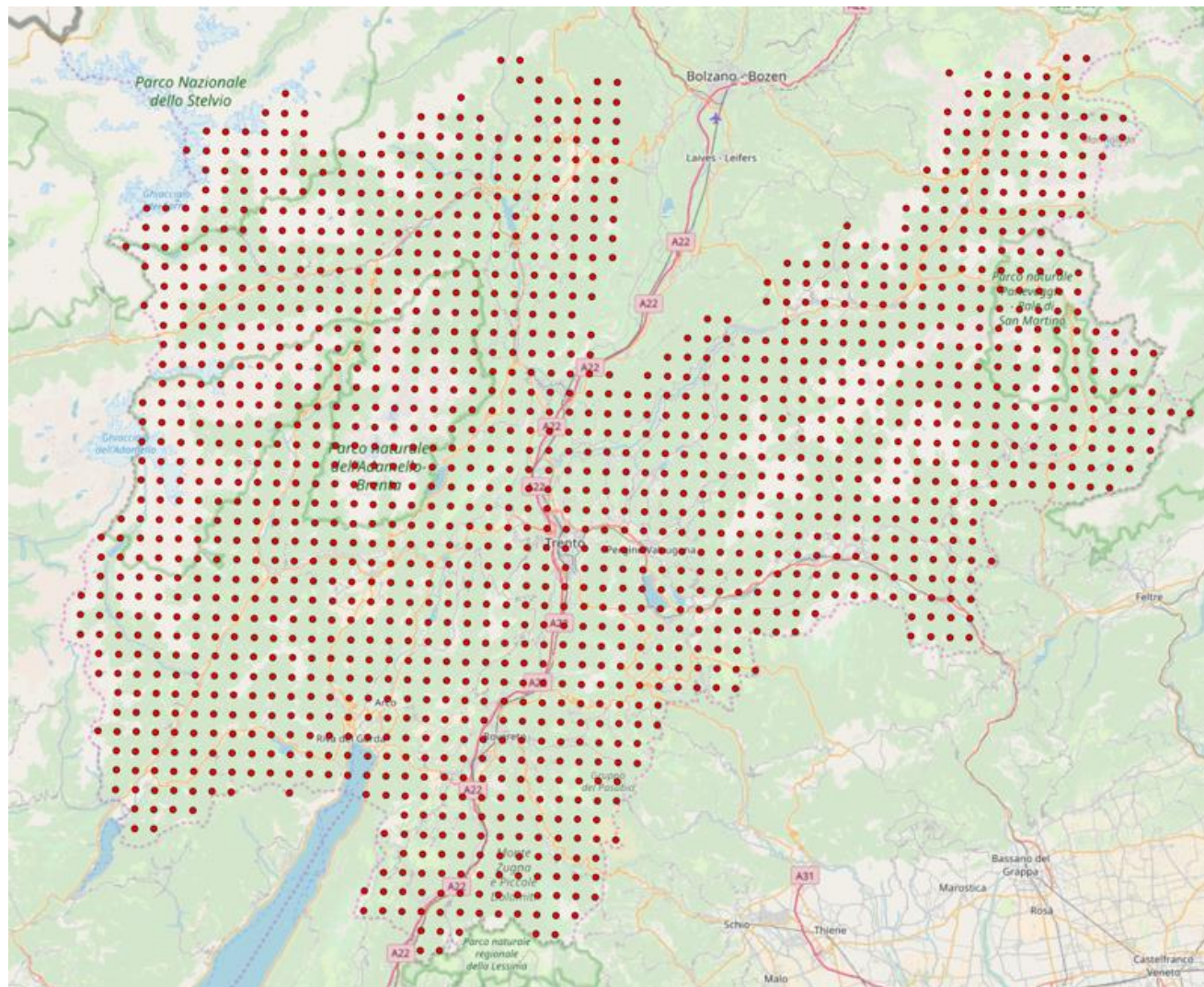


* **Armadio virtuale:** per ogni magazzino consideriamo tutti i prodotti, analizzando per famiglia, principio attivo, categoria di dilavamento/persistenza, prodotto

DSS e trattamenti reali 2017



Meteo utilizzato per la simulazione (fenologia e DSS)



4R Søk etter sted... Avansert søk

Forsiden Italia Trentino og Syd-Tirol Trento

Værvarsel for
Trento, Trentino og Syd-Tirol (Italia)

Oversikt

Time for time
Langtidsvarsel
Været som var
Kart

I dag, tirsdag 18.12.2018

Tid	Varsel	Temp.	Nedbør	Vind
Kl 16–18		8°	0 mm	→ Stille, 0 m/s fra vest
Kl 18–24		6°	0 mm	○ Stille, 0 m/s

- Stazioni Meteotrentino + griglia a 2Km (1541 pt)
- Storico + previsione YR a 2 giorni
- Temperatura, precipitazione e umidità oraria

Risultati Simulazione DSS 2017

DSS

- 49822 trattamenti su tutta stagione,
- 40806 (m=9.2) su finestra rif*
- $\text{DSS} \leq \text{trattamenti reali}$ 99% dei casi

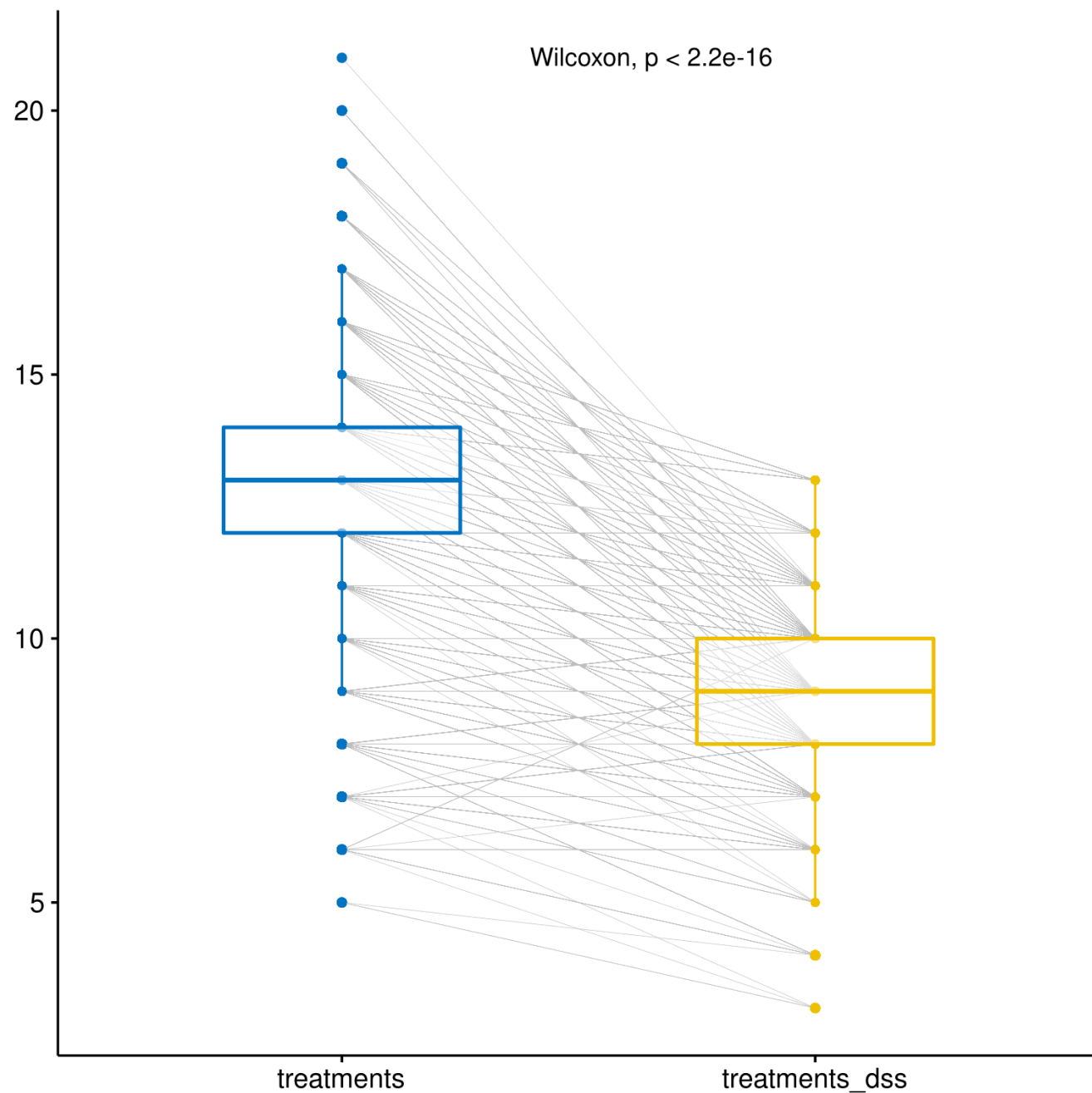
- 29%

2017

- 57476 trattamenti
 - m=12.9
- 70 appezzamenti con $n_{\text{reale}} = n_{\text{DSS}}$
 - Per un campione: 93% con uva sana e matura

(*: intervallo tra primo e ultimo trattamento reale ± 2 giorni)

Dettaglio risultati DSS 2017



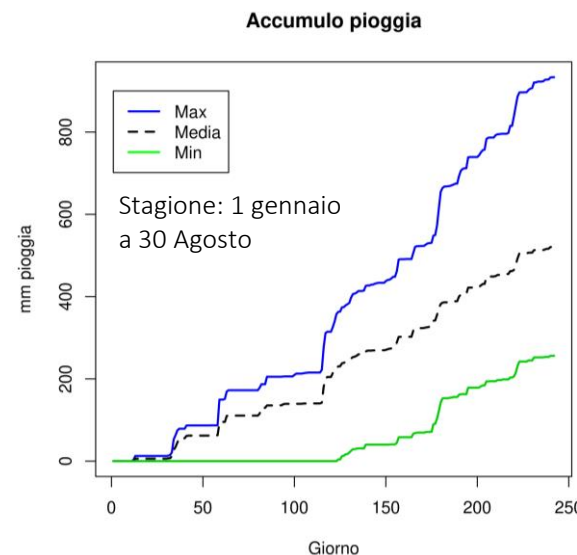
Analisi della simulazione DSS 2017

Ci sono delle variabili comuni negli appezzamenti che trattano molto di più o molto di meno?

Controllo per

- **Geodati:** altitudine, esposizione e pendenza
- **Meteodati:** accumulo di pioggia, umidità, temperature e giorni di pioggia
- **Analisi delle preferenze:**
profiling su preferenze di prodotti

2017 stagione poco piovosa
(20% in meno rispetto a 2016)



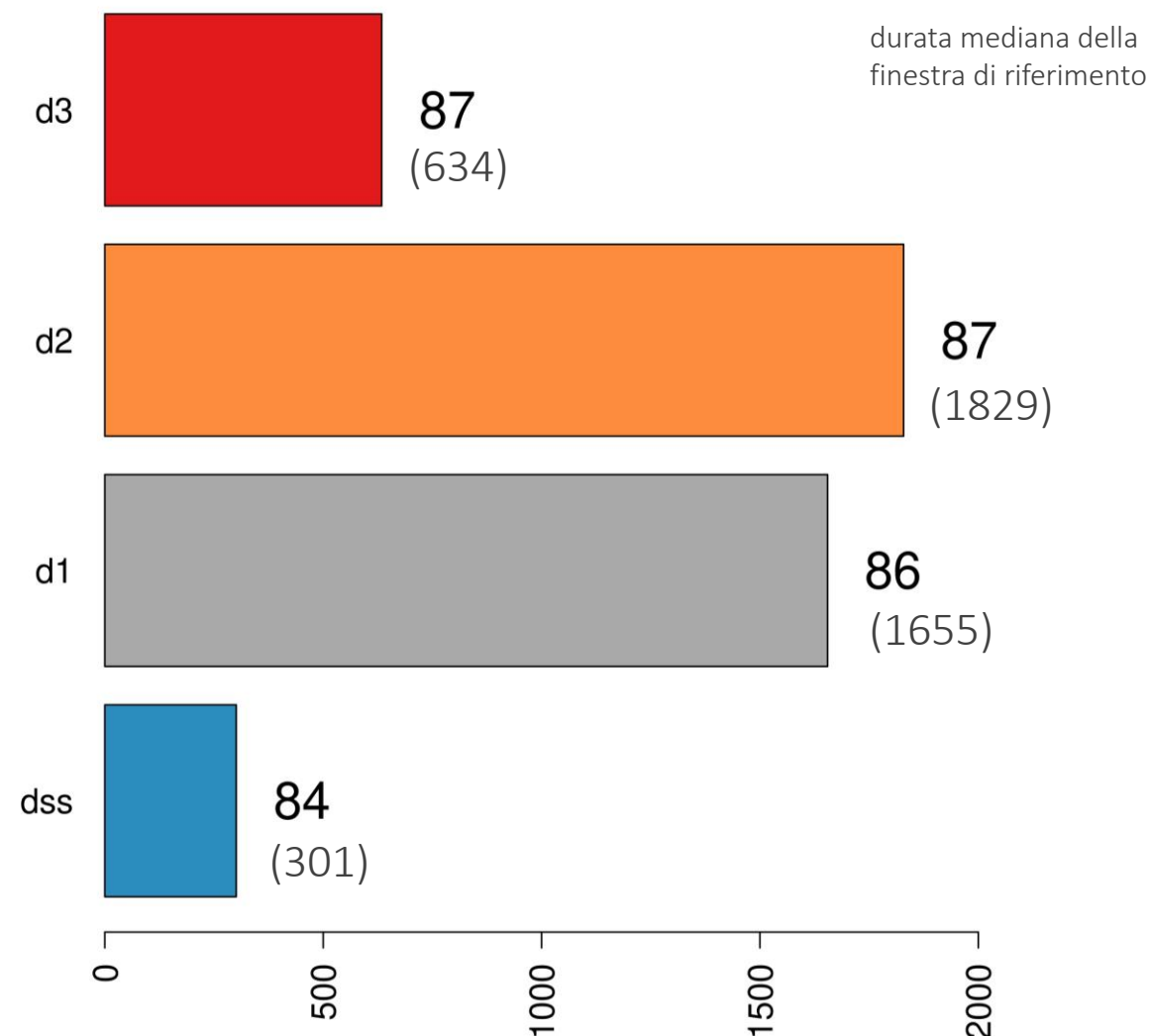
Appezzamento tipo:

- 362m slm, 11% di pendenza, 202° esposizione (SW-W)
- 516 (250 med fin) mm di pioggia,
41% umidità, 96 (56 med fin) giorni piovosi

AI E DEEP LEARNING - Simulazione DSS 2017

Quattro classi di allineamento con DSS
(4419 appezzamenti):

dss: uguali a dss o differenza di ± 1 (301)
d1: 2 o 3 trattamenti in più del dss (1655)
d2: 4 o 5 trattamenti in più del dss (1829)
d3: più di 5 trattamenti in più del dss (634)

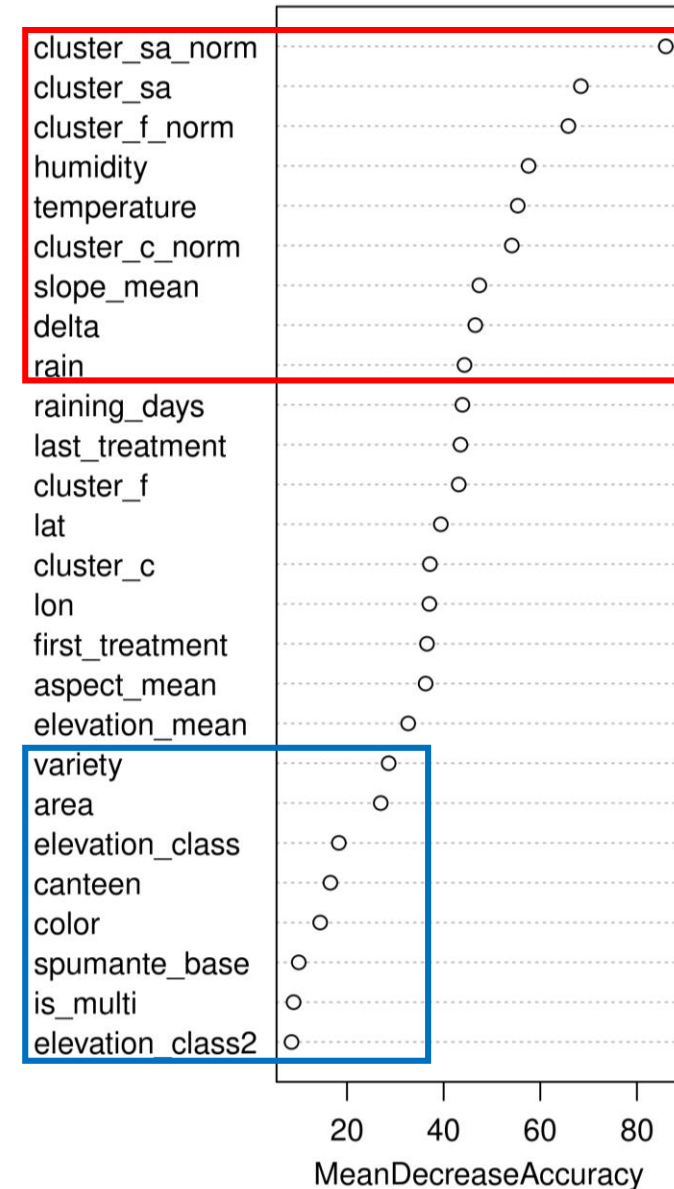


VARIABILI DISCRIMINANTI

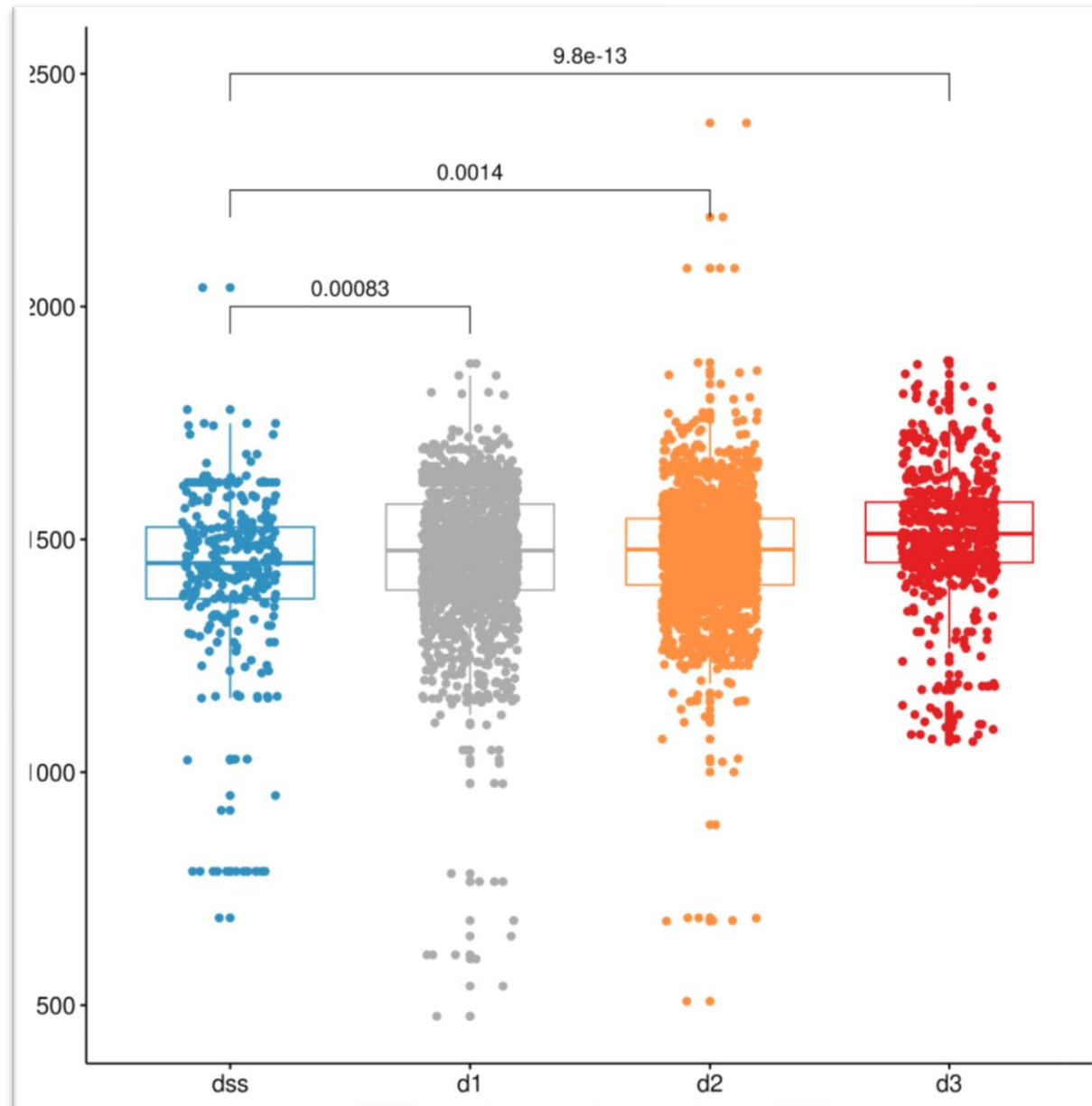
- CLASSIFICATORE Random Forest
- Dati: train su 80%, test su 20%
→ accuratezza MCC test = 0.77 (alta!)

Variabili più importanti: **profiling SA, umidità, temperatura, finestra operativa**

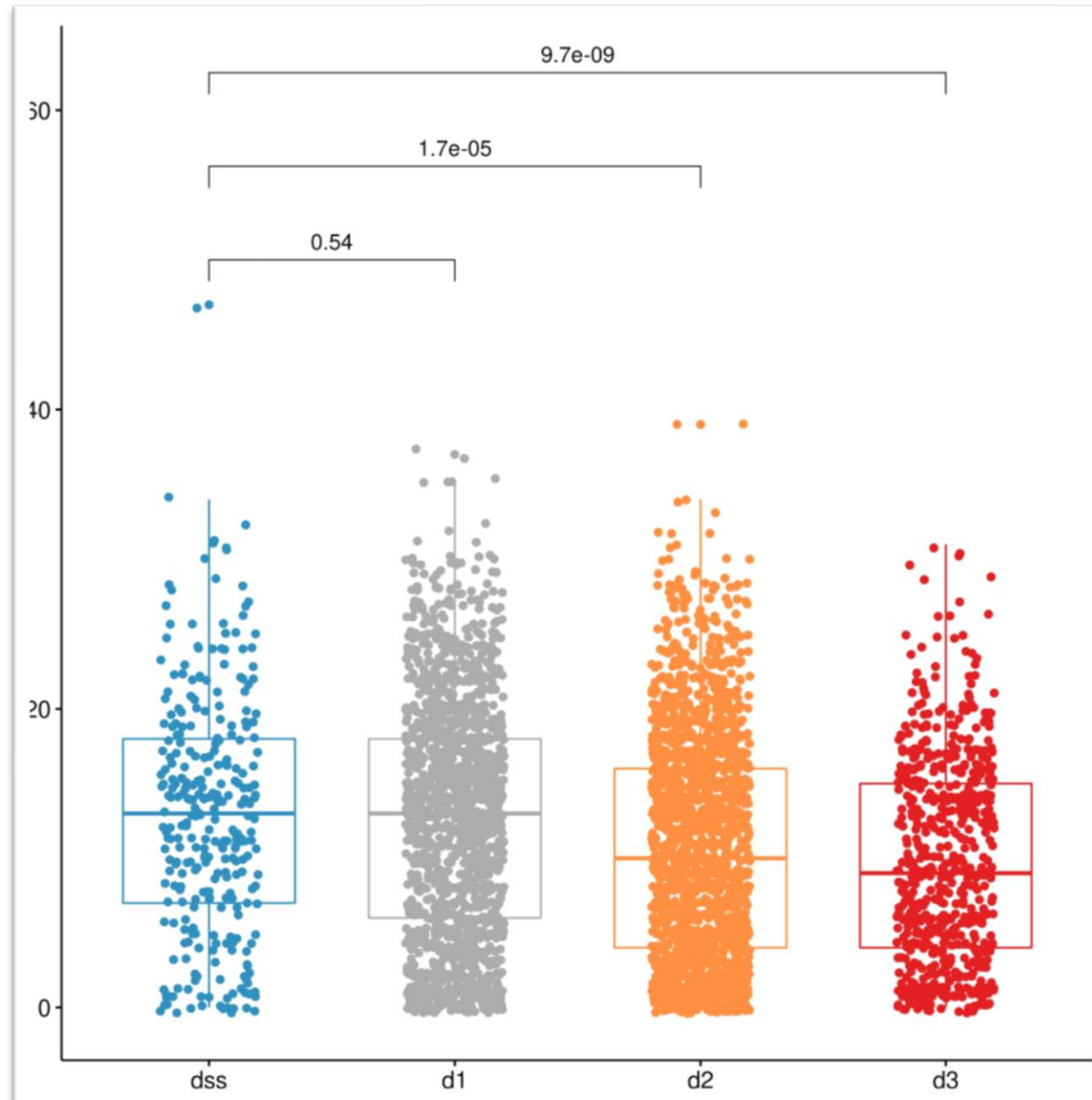
Variabili meno importanti: **cantina, caratteristiche uva**



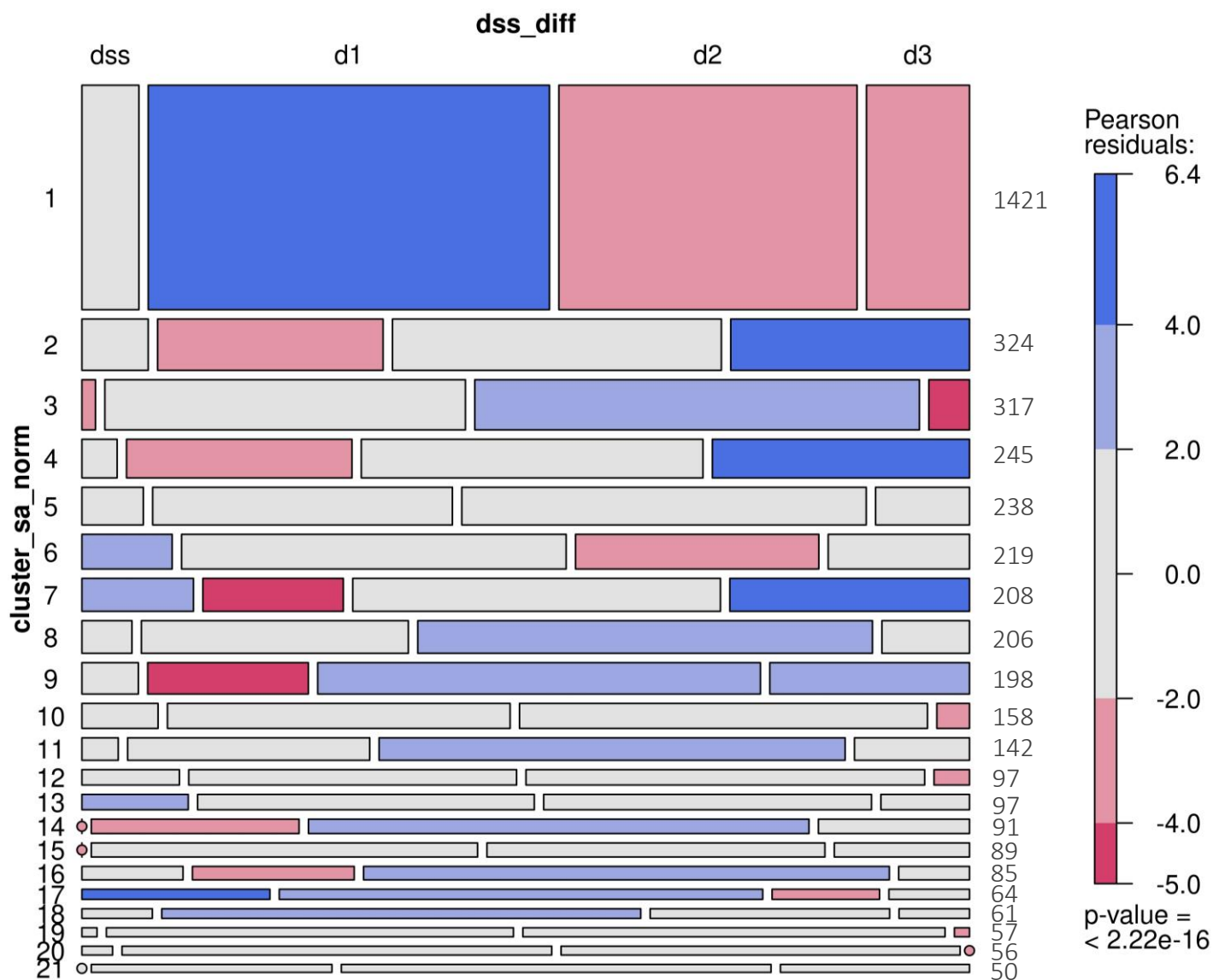
ANALISI CLASSIFICATORE – Accumulo temperatura



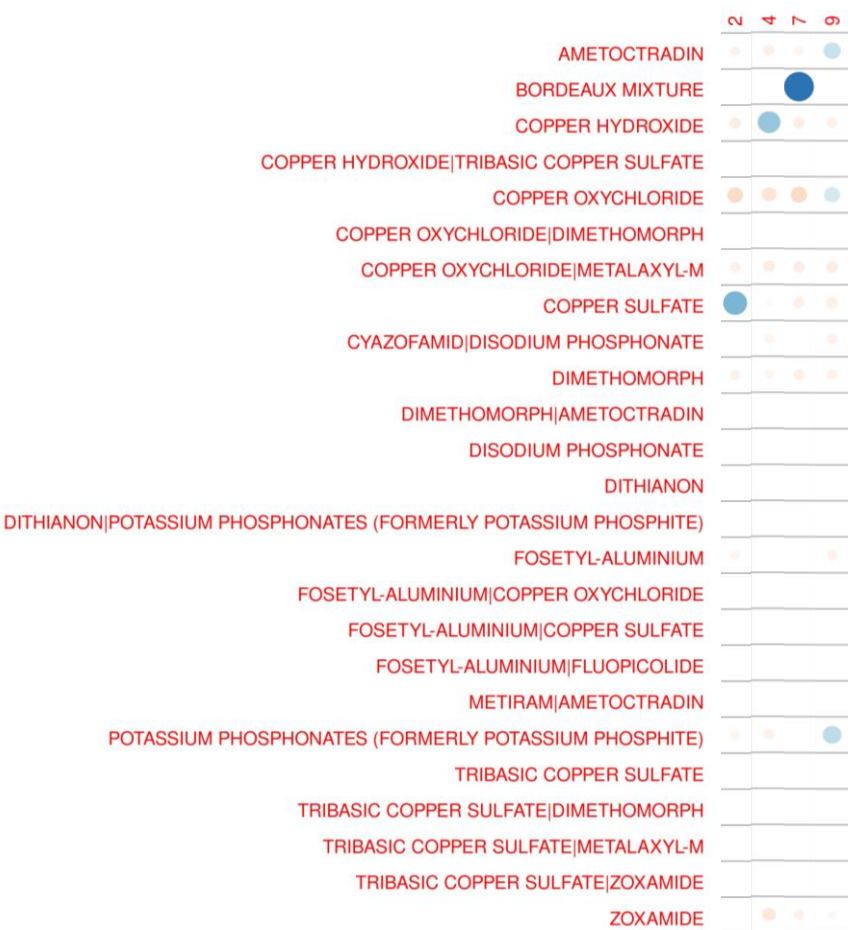
ANALISI CLASSIFICATORE – Pendenza media



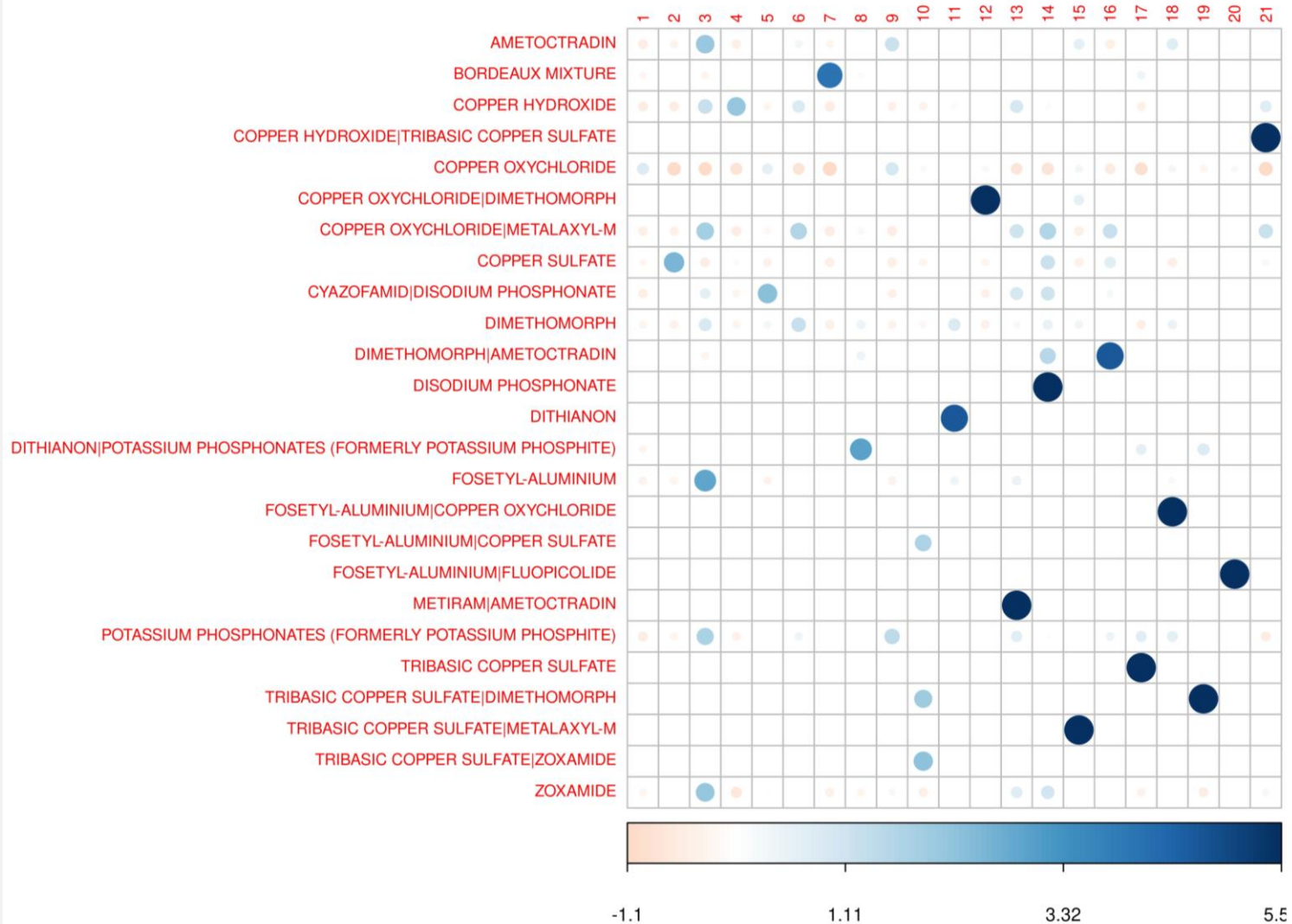
AI E DEEP LEARNING – Profiling prodotti



- Gruppi 2, 4, 7, 9
numerosità >> in classe d3



SA_NORM – Profiling prodotti

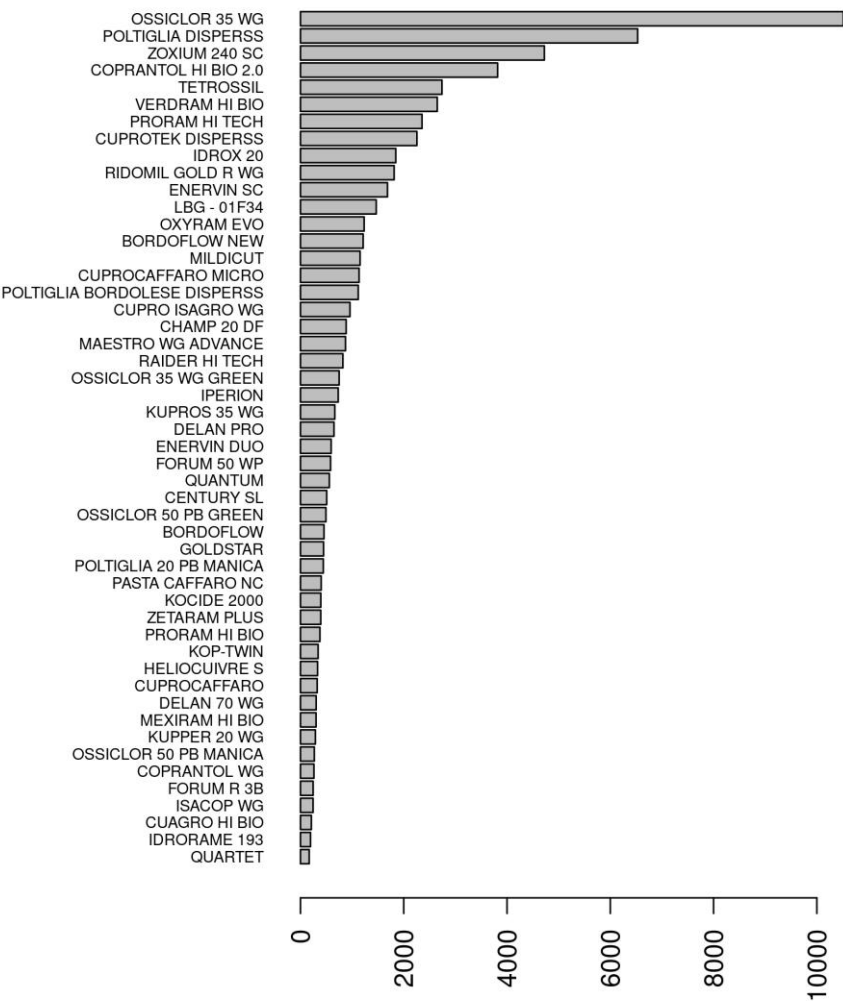


Gruppi di appezzamenti che utilizzano pattern simili di prodotti (sostanze attive)

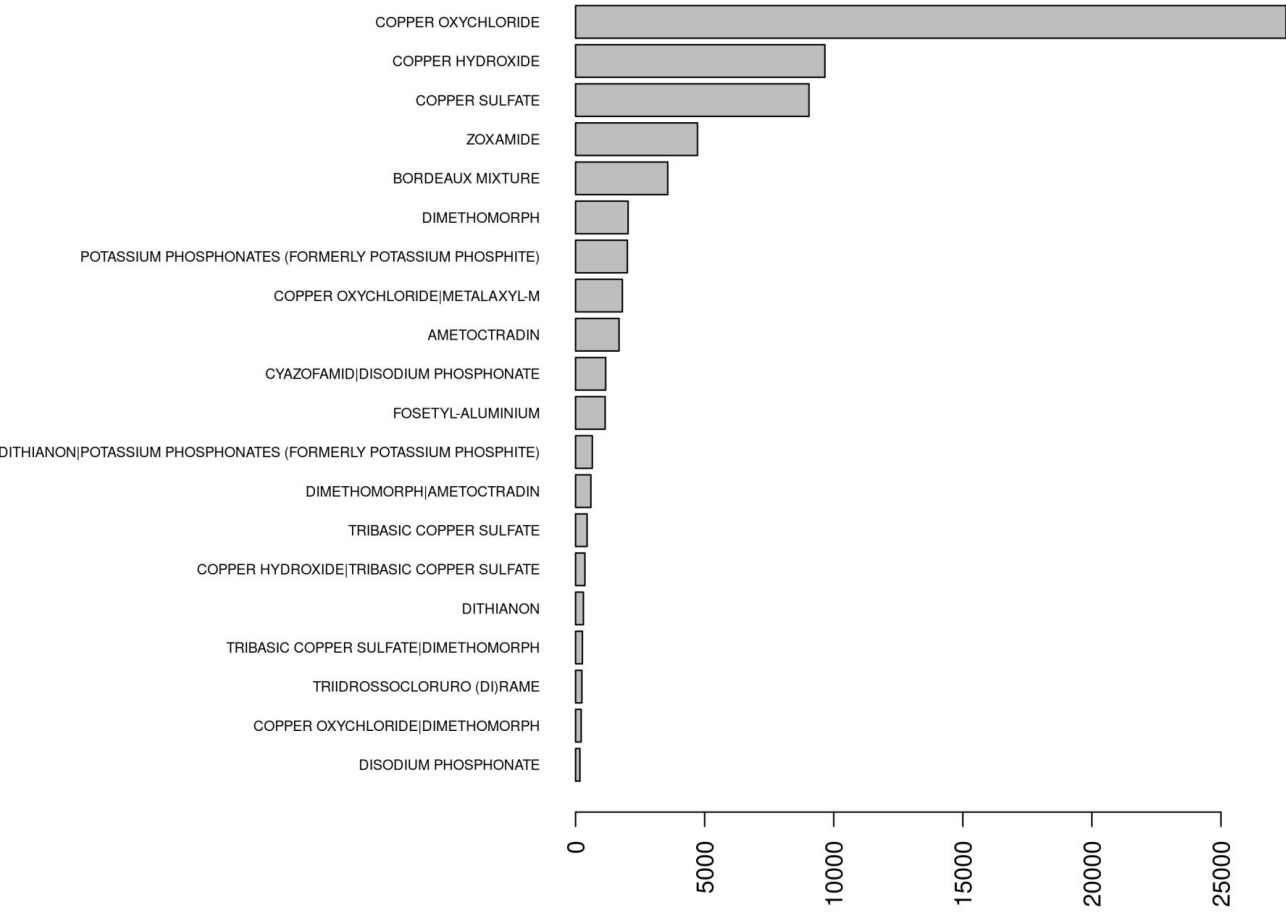
Profiling prodotti nel magazzino virtuale

DB prodotti: arricchimento DB ministeriale con limiti utilizzo e restrizioni (MPA Solution)

TOP 50 Prodotti

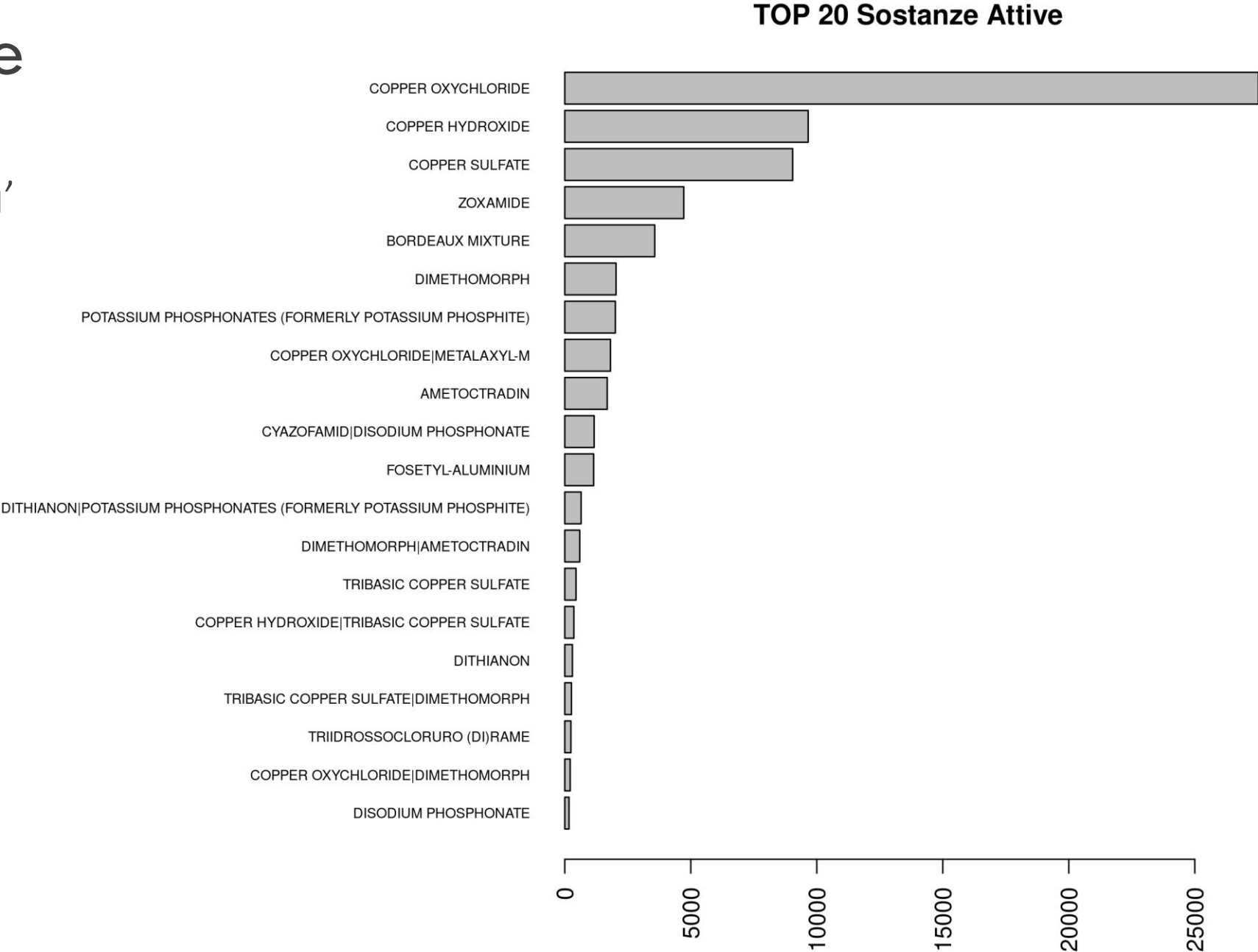


TOP 20 Sostanze Attive



Sostanze attive

Nei 57k trattamenti,
le 3 sostanze attive piu'
utilizzate sono rameici
e coprono il 67%

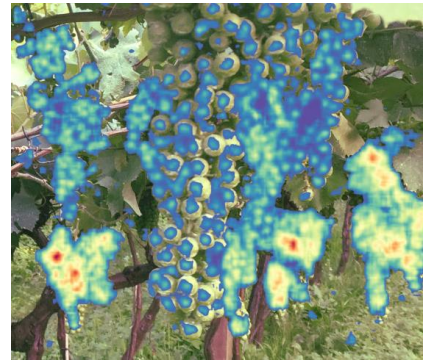


Progetti 2019: AI per Agricoltura Digitale

FELLOWSHIP “Deep Learning per il monitoraggio della qualità del raccolto”

- Sviluppo di modelli predittivi da sensori a basso costo
- Smart Farming: cloud e edge computing

GARR: rete nazionale a banda ultralarga per l'istruzione e la ricerca



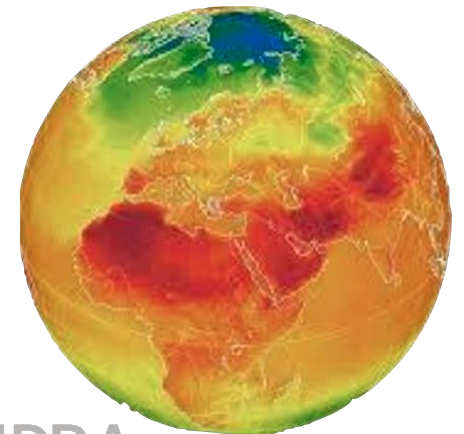
AI for earth

A \$50 million commitment with an expanded strategic approach

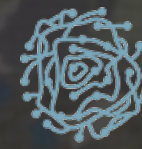


AWARD a FBK “Deep Learning per previsione degli impatti crop-specific delle ondate di calore ”

- Onde di calore in aumento a causa dei cambiamenti climatici
- Effetti disastrosi sull'ambiente e la crescita delle colture
- Goal: predire rischi e danni con alta precisione



AI per Agricoltura Digitale



MPBA

PREDICTIVE MODELS
FOR BIOMEDICINE & ENVIRONMENT



Cesare Furlanello
(Responsabile Unità)



Marco Cristoforetti
(Ricercatore)



Gabriele Franch
(Dottorando)



Azra Alikadić
(Ricercatore)



Andrea Gobbi
(Ricercatore)



Luca Coviello
(Fellowship)



Giulia Bruscagin
(Project Manager)



Francesco Ballerin
(Tecnologo Junior)



Andrea Zanin
(Tecnologo Junior)

Ringraziamenti a: Giuseppe Jurman e Valerio Maggio