

FONDO SOCIALE EUROPEO IN SINERGIA CON IL FONDO EUROPEO DI SVILUPPO
REGIONALE
POR 2014-2020 – OBIETTIVO "INVESTIMENTI A FAVORE DELLA CRESCITA E
DELL'OCCUPAZIONE"
STRATEGIE REGIONALI PER IL SISTEMA UNIVERSITARIO
"INNOVAZIONE E RICERCA PER UN VENETO PIÙ COMPETITIVO
ASSEGNI DI RICERCA - ANNO 2019"
DGR NR. 1463 DEL 08/10/2019



UNIVERSITÀ
di **VERONA**

Digital Transformation Labs to Observation, Business, Artefacts and Spread – RACCOLTA INTELLIGENTE DATI E ANALISI STATISTICHE VITIVINICOLE

COD. ENTE 1695 UNIVERSITA' DEGLI STUDI DI VERONA
ASSE OCCUPABILITA' - DGR NR. 1463 DEL 08/10/2019

Codice Progetto	2105-0068-1463-2019
Referente progetto	Franco Fummi
Assegnista	Enrico Carlo Agrillo
Dipartimento	Informatica

INDICE

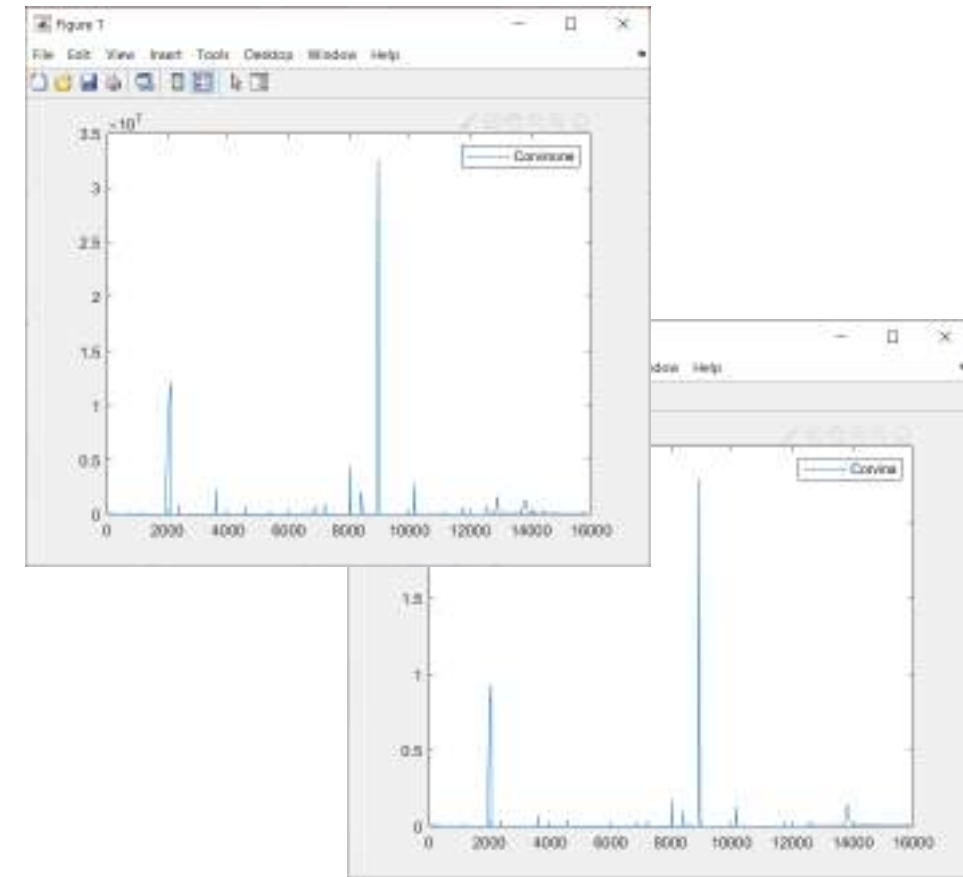
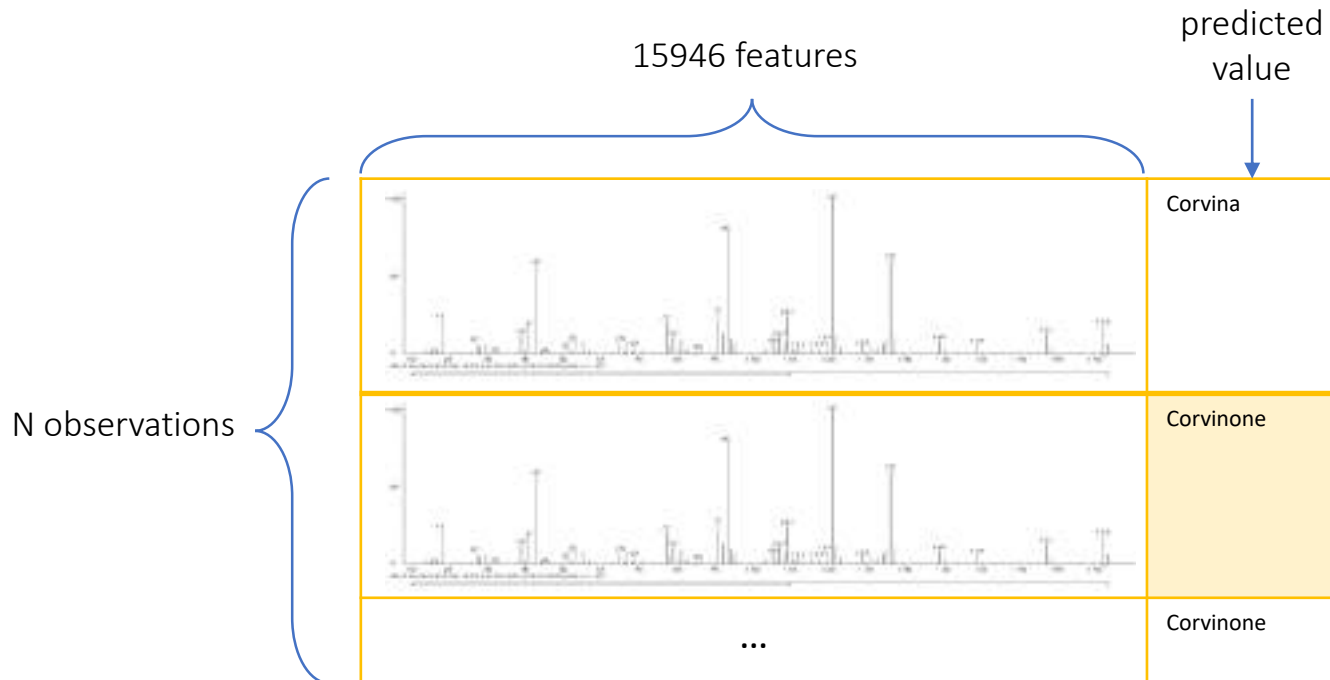
- Obiettivo della ricerca
- Analisi dati e riconoscimento Corvina e Corvinone
- Feature selection e risultati
- Docker, Grafana ed InfluxDB
- Realizzazione dashboard e plugin personalizzato
- Studio degli analizzatori
- Soluzione per interfacciamento sensori con architettura dati

Obiettivo della ricerca

- Effettuare raccolta di dati ambientali per monitorare la qualità lavorativa
- Costruire e realizzare un'architettura composta da sensori enologici, gateway e cloud che permetta la raccolta dei dati sensoriali e chimici
- Analisi dati tramite l'utilizzo di tecniche di machine learning.

Analisi dati e riconoscimento Corvina e Corvinone

- 40 cromatografie di due tipi di vino
 - Corvina
 - Corvinone
- Una cromatografia contiene 15946 feature

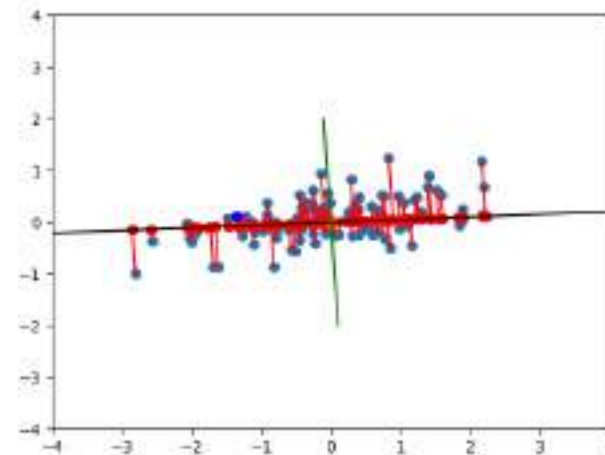


Analisi dati e riconoscimento Corvina e Corvinone

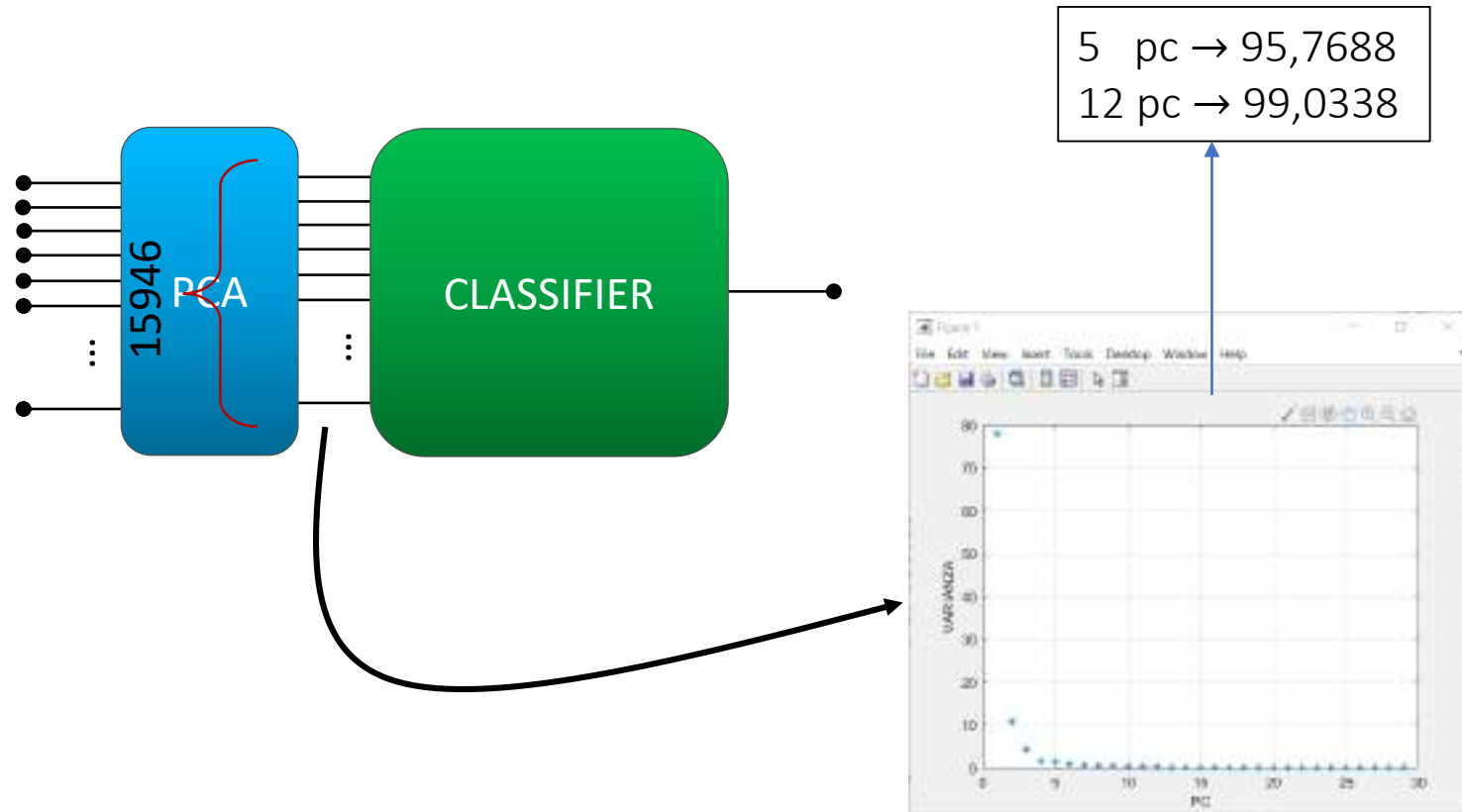
- Allenamento del dataset con 20/40 campioni con tutte le feature
 - Classifier: Quadratic SVM
 - Prediction phase: 10/20 error
- Allenamento del dataset con 30/40 campioni con tutte le feature
 - Classifier: Quadratic SVM
 - Prediction phase: 1/10 error
- Allenamento del dataset con 20/40 campioni con 10000/15946 feature
 - Classifier: linear SVM
 - Prediction phase: 1/4 error

Analisi dati e riconoscimento Corvina e Corvinone

- Problema: troppe feature per campione
- Possibile soluzione:
 - utilizzare PCA (Principal Component Analysis)
- Tecnica di semplificazione dei dati
 - Obiettivo: ridurre il numero di variabili che descrivono un dato ad un più piccolo numero di variabili latent, limitando la perdita di informazioni il più possibile



Analisi dati e riconoscimento Corvina e Corvinone



Analisi dati e riconoscimento Corvina e Corvinone

- Allenamento del dataset con 30/40 campioni e 4pc (94.412% information)
 - Classifier: Subspace KNN
 - Prediction phase: 7/10 error
- Allenamento del dataset con 30/40 campioni e 5pc (95.768% information)
 - Classifier: linear discriminant
 - Prediction phase: 3/10 error
- Allenamento del dataset con 30/40 campioni e 6pc (96.687% information)
 - Classifier: Quadratic discriminant
 - Prediction phase: 5/10 error
- Allenamento del dataset con 30/40 campioni e 12pc (99.033% information)
 - Classifier: Subspace KNN
 - Prediction phase: 5/10 error

Feature selection e risultati

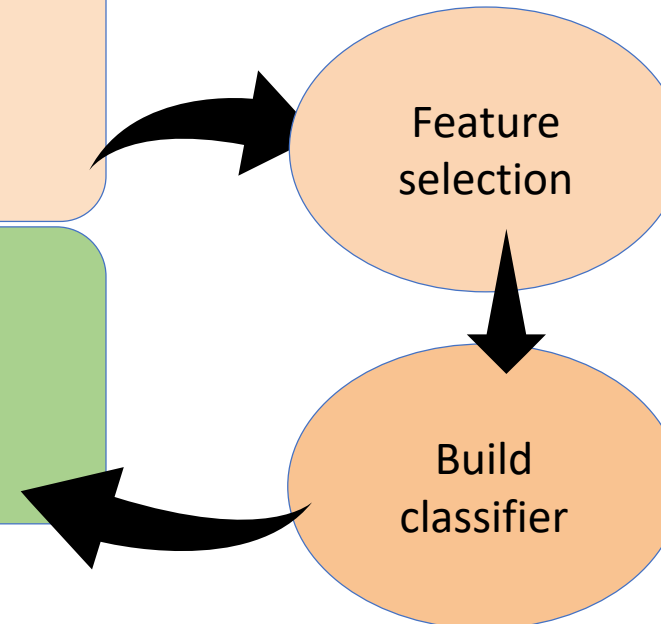
- PCA non migliora le performance
 - Il processo è più rapido ma le predizioni sono peggiori
- Obiettivo: Costruire un classificatore che, dato in input la cromatografia di un vino, dica se è Corvina o Corvinone

- Training dataset:

- 40 cromatografie di due tipi di vino
 - 20 di Corvina
 - 20 di Corvinone

- Testing dataset:

- 30 cromatografie di due tipi di vino
 - 15 di Corvina
 - 15 di Corvinone

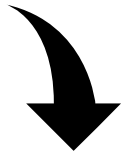


Feature selection e risultati

- Ogni processo vengono selezionate le 10 feature più significative
- Ripetuto per 500 processi

“Infinite Feature Selection” method[1]

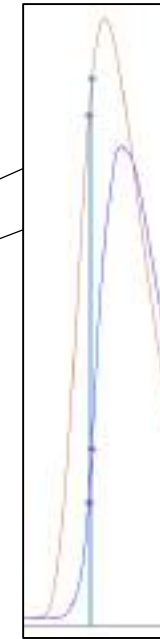
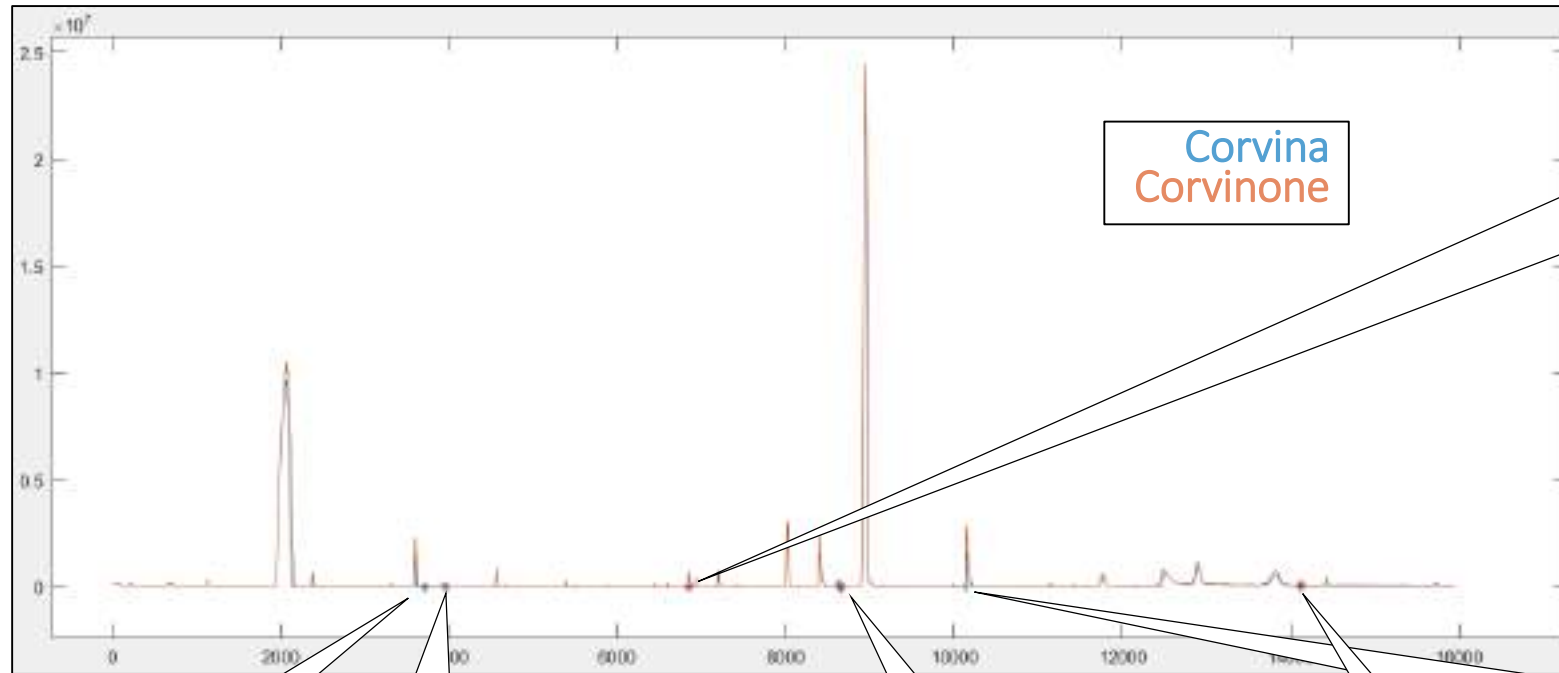
Solves the feature selection problem through a graph where the nodes represent the features and the arcs represent the relations between them. These relationships are defined in terms of correlation and standard deviation combined by an alpha parameter



	1	2	3	4	5	6	7	8	9	10
1	3955	3956	3957	3954	3958	3953	3952	3951	3950	3959
2	8639	8638	8637	8640	8635	8634	8636	8633	8632	8641
3	8639	8640	8638	8637	8636	8641	8633	8642	8634	8643
4	3952	3953	3951	3950	3954	3949	3955	3948	3947	3946
5	8639	8638	8640	8637	8636	8641	8633	8642	8643	8634
6	8639	8638	8640	8637	8636	8641	8633	3956	3955	8642
7	8639	8638	8637	8636	8640	8635	8641	8634	8642	8633
8	8639	8640	8641	8638	8642	8645	8644	8643	8637	8646
9	8638	8639	8637	8636	8640	8635	8641	8634	8642	3957
10	3955	3956	3954	3957	3953	3958	3952	3951	3950	3949
11	3956	3957	3955	3954	3958	3953	3952	3951	3959	3950
12	3953	3952	3954	3951	3950	3955	3949	3948	3956	3947
13	3957	3956	3955	3954	3953	3958	3952	3951	3950	3959
14	8639	8646	3957	8640	8645	3956	8638	3958	8641	8647
15	8639	8638	8637	8640	8636	8635	8641	8634	8642	8643
16	3955	3954	3953	3956	3952	3957	3951	3950	3958	3949
17	8639	8637	8638	8636	8640	8635	8641	8634	8642	8633
18	8639	8638	8637	8636	8635	8640	8634	8633	8641	8632
19	8645	8639	8640	8646	8641	8644	8654	8655	8642	8647
20	3957	3955	3954	3956	3958	3953	3952	3951	3950	3949
21	3954	3955	3953	3956	3952	3957	3951	3958	3950	3949
22	8639	8640	8641	3957	8638	3950	3955	3958	8642	3954
23	8639	8640	8638	8641	8637	8636	8640	8635	8643	8634

[1] Roffo Giorgio, Simone Melzi, and Marco Cristani. "Infinite feature selection". *Proceedings of the IEEE International Conference on Computer Vision*. 2015.

Feature selection e risultati



Molecola:
acido 3-metilbutanoico

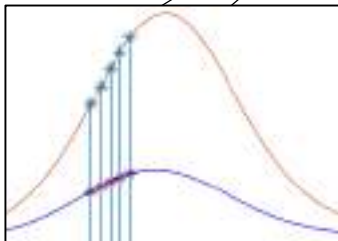
Odore: rancido,
butirrico, sgradevole

precursore dell'estere etil 3-
metilbutanoato che presenta
aromi fruttati

Molecola:
acido ottanoico

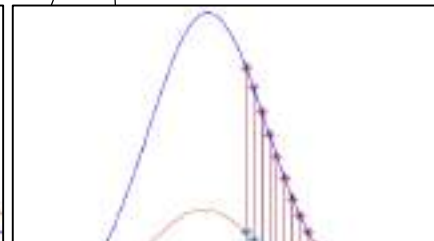
Odore: rancido,
butirrico, sgradevole

precursore dell'estere etil
ottanoato con odore
piacevole, fruttato



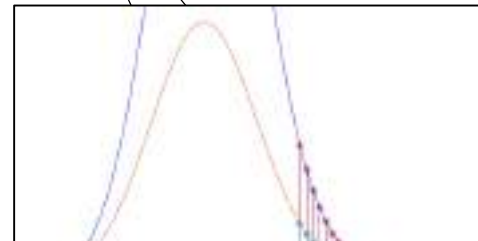
Molecola:
cis-3-esen-1-olo

Sentori: erba
tagliata

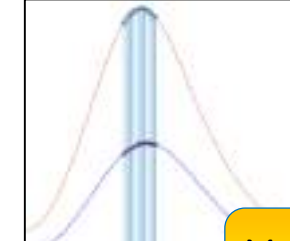


Molecola:
trans-3-esen-1-olo

Sentori: erba
tagliata



Molecola:
alcol benzilico



Molecola:
etil vanillato

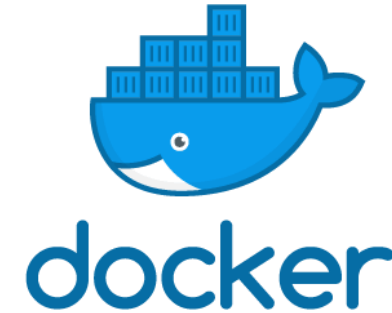
Odore: aroma poco
intenso, speziato,
vanigliato

Feature selection e risultati

- I risultati dimostrano che le molecole selezionate dall'algoritmo corrispondono a quelle utilizzate durante le analisi
- In particolare il Corvina e il Corvinone sono distinti per i loro contenuti di cis e trans 3-esenolo ed etil vanillato

Docker, Grafana ed InfluxDB

- Docker come servizio di containerizzazione
- InfluxDB come database NoSQL
- Grafana per la visualizzazione dei dati tramite dashboard



Realizzazione dashboard e plugin personalizzato

- Necessario sviluppare un plugin personalizzato
- Possibilità di visualizzare i dati di monitoraggio ambientale
- I dati sono contenuti all'interno di PanelData
 - Series di tipo DataFrame contiene gli endpoint
 - Fields di tipo Field contiene il tipo di dato (temperatura, umidità, ...)
 - Values di tipo Vector contiene invece i valori da visualizzare

```
export interface DataFrame extends QueryResultBase {  
  name?: string;  
  fields: Field[];  
  length: number;  
}
```

```
export interface PanelData {  
  /** State of the data (loading, done, state: LoadingState;  
  /** Contains data frames with field ov series: DataFrame[];  
  /** A list of annotation items */ annotations?: DataFrame[];  
  /** Request contains the queries and p request?: DataQueryRequest;  
  /** Timing measurements */ timings?: DataQueryTimings;  
  /** Any query errors */ error?: DataQueryError;  
  /** Contains the range from the request timeRange: TimeRange;  
}
```

```
export interface Field<T = any, V = Vector<T>> {  
  /**  
   * Name of the field (column)  
   */  
  name: string;  
  /**  
   * Field value type (string, number, etc)  
   */  
  type: FieldType;  
  /**  
   * Meta info about how field and how to display it  
   */  
  config: FieldConfig;  
  values: V;  
  labels?: Labels;  
  /**  
   * Cached values with appropriate display and id values  
   */  
  state?: FieldState | null;  
  /**  
   * Convert text to the field value  
   */  
  parse?: (value: any) => T;  
  /**  
   * Convert a value for display  
   */  
  display?: DisplayProcessor;  
  /**  
   * Get value data links with variables interpolated  
   */  
  getlinks?: (config: ValueLinkConfig) => Array<LinkModel<Field>>;  
}
```

Realizzazione dashboard e plugin personalizzato

- Il codice contiene una sezione html per la visualizzazione ed una per l'estrazione dei dati e la logica



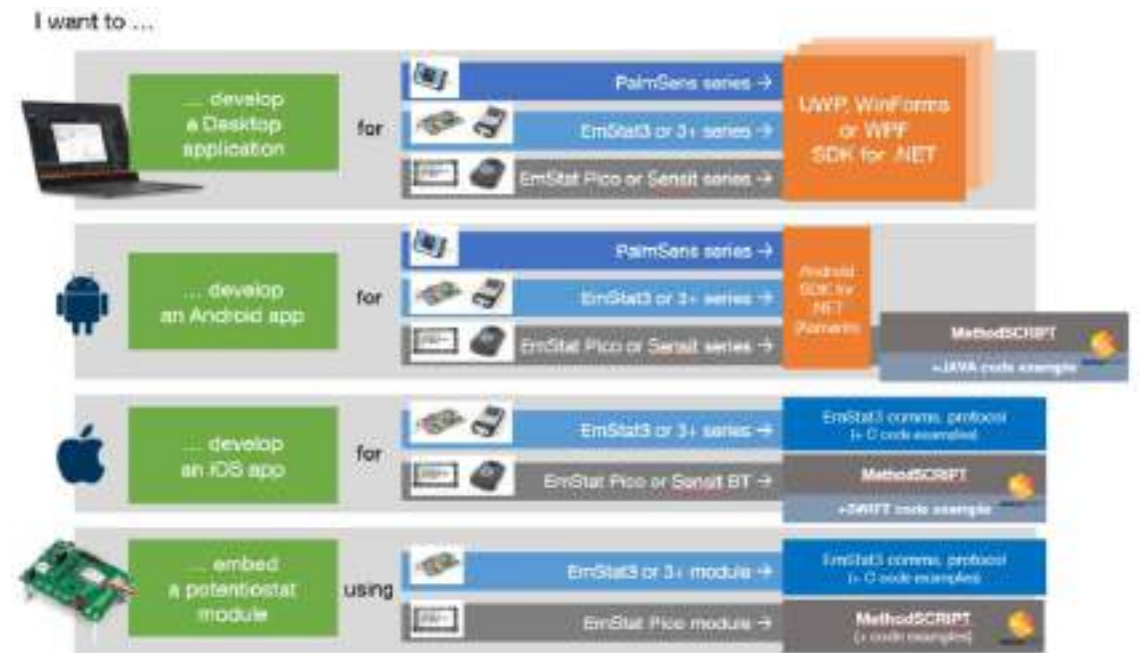
Studio degli analizzatori

- Obiettivo: comunicazione tra gli analizzatori e l'architettura dati del laboratorio ICE
 - BOX-IO viene già utilizzato con diversi sensori
 - La comunicazione può avvenire con o senza cavo
- Una volta connessi gli analizzatori a BOX-IO, i dati possono essere inviati all'architettura dati del laboratorio ICE



PalmSes4

- Sviluppare una comunicazione automatica con un sistema embedded non è possibile
 - Questo tipo di comunicazione è disponibile solo per il modulo EmStat
- Per la comunicazione del PalmSens4 possiamo accedere a degli SDK .NET
 - Possiamo sviluppare una soluzione software UWP che comunichi con l'analizzatore



Nomasens

- Nomasense Polyscan P200, Color P100 and P6000 Oxygen Analyzer
- La **comunicazione richiede un software specifico** proprietario e sviluppato da una compagnia esterna
 - Le analisi dei risultati sono inviati ad un server del **cloud “Nomascan”**
- È possibile **scaricare i file CVS** cche contengono le informazioni e **sviluppare una soluzione software**
 - Per la manipolazione dei dati ed il popolamento automatico dell’architettura dati

```
Exported with Datamanager, Version=1.0.9.0, Culture=neutral, PublicKeyToken=null V1.0.9.0
Data;Time;User;SensorID;SensorName;delta_t;Unit;Value;Unit;Phase;Unit;Amplitude;Unit;Temp;Unit;patm;Unit;Alt;Loww
30-May-2016;19:44:21;default ;189276 ;COX ;0.00;min;21.386;MO2;27.083;deg;1486;uV;20.621;°C;1015.038;mbars;
30-May-2016;19:44:24;default ;189276 ;COX ;0.05;min;21.346;MO2;26.982;deg;1410;uV;20.621;°C;1015.150;mbars;
30-May-2016;19:44:27;default ;189276 ;COX ;0.10;min;21.208;MO2;27.058;deg;1410;uV;20.621;°C;1015.198;mbars;
30-May-2016;19:44:30;default ;189276 ;COX ;0.15;min;21.416;MO2;26.947;deg;1411;uV;20.621;°C;1015.218;mbars;
30-May-2016;19:44:33;default ;189276 ;COX ;0.20;min;21.183;MO2;27.062;deg;1407;uV;20.621;°C;1015.250;mbars;
30-May-2016;19:44:50;default ;189276 ;COX ;0.48;min;21.313;MO2;27.088;deg;2571;uV;20.621;°C;1014.978;mbars;
30-May-2016;19:44:52;default ;189276 ;COX ;0.52;min;21.182;MO2;27.063;deg;2572;uV;20.621;°C;1015.210;mbars;
30-May-2016;19:44:55;default ;189276 ;COX ;0.57;min;21.192;MO2;27.058;deg;2574;uV;20.621;°C;1015.198;mbars;
30-May-2016;19:44:58;default ;189276 ;COX ;0.62;min;21.326;MO2;26.991;deg;2577;uV;20.621;°C;1015.250;mbars;
30-May-2016;19:45:01;default ;189276 ;COX ;0.67;min;21.306;MO2;27.001;deg;2578;uV;20.621;°C;1015.240;mbars;
30-May-2016;19:45:04;default ;189276 ;COX ;0.72;min;21.139;MO2;27.085;deg;2574;uV;20.621;°C;1015.158;mbars;
30-May-2016;19:48:40;default ;189276 ;COX ;4.32;min;20.548;MO2;27.385;deg;3803;uV;20.621;°C;1015.140;mbars;
```

```
SourceDeviceSerialNumber,SourceDeviceFamiliarName,CreatorLoginName,DataTime,isSeries,isUsingDefaultTitle,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-18 13:14:48","1178A3A9-5368-44CA-EA77-2898DD67A525,
"7b1468168f8ddea","Prototype Labo","EBtest","2020-08-20 18:29:07","1178A3A9-5368-44CA-EA77-2898DD67A525,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-03 18:17:12","8FF3573B-9E88-4D52-98E6-24A158F238C8,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-03 18:14:23","8FF3573B-9E88-4D52-98E6-24A158F238C8,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-03 18:11:09","8FF3573B-9E88-4D52-98E6-24A158F238C8,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-02 10:49:12","8FF3573B-9E88-4D52-98E6-24A158F238C8,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-02 10:47:12","8FF3573B-9E88-4D52-98E6-24A158F238C8,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-02 10:45:24","8FF3573B-9E88-4D52-98E6-24A158F238C8,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-02 10:43:50","8FF3573B-9E88-4D52-98E6-24A158F238C8,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-02 10:41:57","8FF3573B-9E88-4D52-98E6-24A158F238C8,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-02 10:33:21","8FF3573B-9E88-4D52-98E6-24A158F238C8,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-02 10:31:38","8FF3573B-9E88-4D52-98E6-24A158F238C8,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-02 10:30:07","8FF3573B-9E88-4D52-98E6-24A158F238C8,
"7b1468168f8ddea","Prototype Labo","EBtest","2021-02-02 10:28:36","8FF3573B-9E88-4D52-98E6-24A158F238C8,
```

Bevscan

- Non possono essere fornite **SDK**
- I dati sono salvati in una **specific cartella** sul laptop connesso
- I file .bvs sono un **file di testo XML delimitato da virgole** strutturato come segue*:
 - 36 righe che mostrano filename, parametri interni, risultati SimCal, risultati delle predizioni etc
 - 601 campi dati con letture dello spettro per lunghezze d'onda che fanno da 450nm a 1050nm con un 1nm di incremento

*Esempio di un BevScan Beverage Analyser BS01

C:\BevScan	The master directory in the root directory of the Touchscreen.
C:\BevScan\Configuration	A critical directory containing the operational data for BevScan - do not change or alter any data here.
C:\BevScan\Data	The master directory for all measurement results in BevScan.
C:\BevScan\Data\Scan (C:\BevScan\Data\Scan\20100701)	The directory containing all Scan Mode spectral results in dated subdirectories.
C:\BevScan\Data\SimCal (C:\BevScan\Data\SimCal\001\20100701)	A hierarchal directory containing all SimCal Mode spectral results in numbered subdirectories with subordinate dated subdirectories.
C:\BevScan\Data\Prediction (C:\BevScan\Data\Prediction\001\20100701)	A hierarchal directory containing all Prediction Mode spectral results in numbered subdirectories with subordinate dated subdirectories.
C:\BevScan\Data\Import	The directory where all data for importing SimCals etc should be placed.
C:\BevScan\Data\Import\SimCal Spectra	The directory in which to place .bvs format files for creating a new SimCal or adding to a current SimCal. Accessed from the [Import Spectra] button in the SimCal - Create/Add Samples page.
C:\BevScan\Data\Import\External Spectra	The directory in which to place .bvs format files for comparing within a SimCal. Accessed from the [Process External Spectra] button in the SimCal Mode page.
C:\BevScan\Data\Export	The directory where all exported data such as SimCals, configurations, backups etc are stored.
C:\BevScan\Data\Export\Processed Spectra	The directory in which the updated .bvs format files are placed once they have been imported for comparison against a current SimCal. Accessed from the [Recall Processed Spectra] button in the SimCal Mode page.

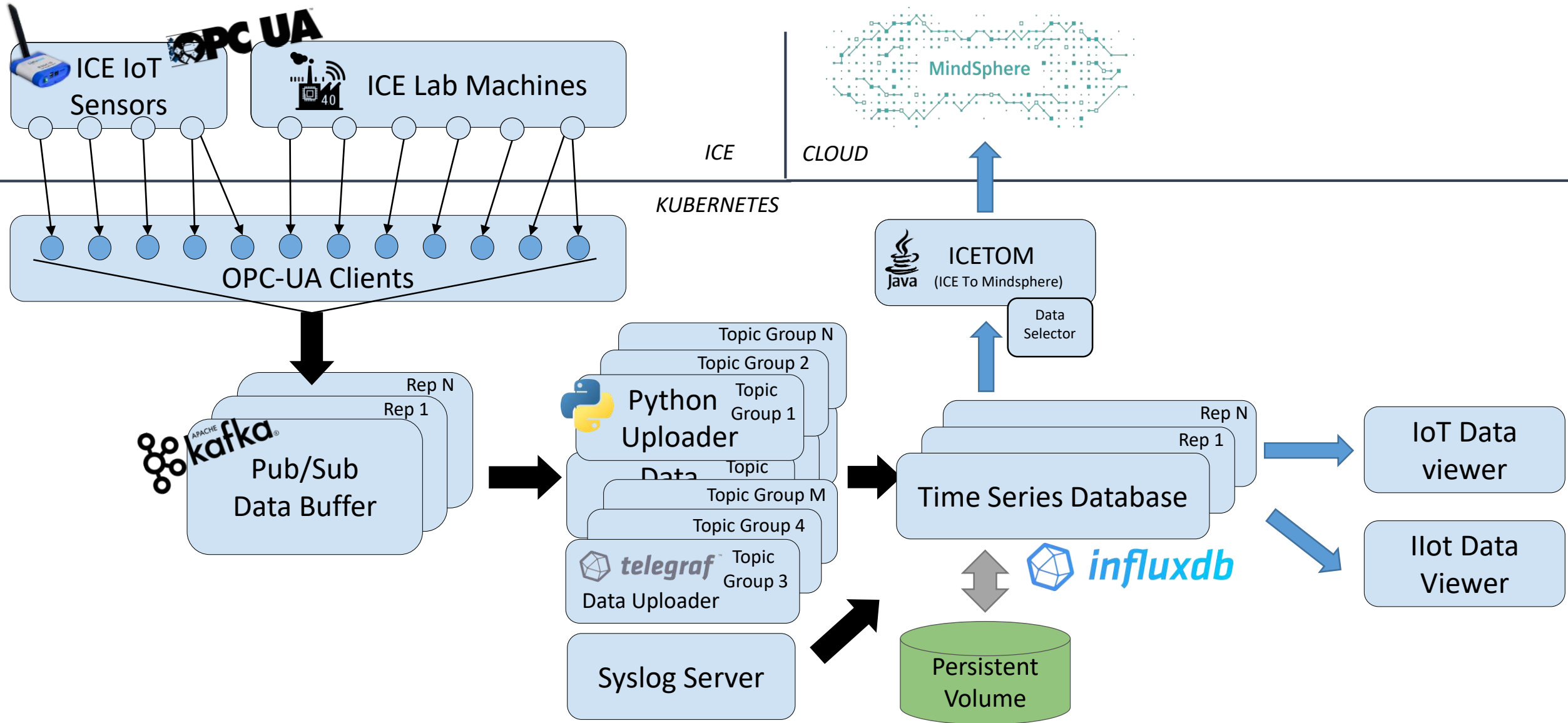
Y15

- **Y15** comunica scrivendo e leggendo in **specifiche cartelle**
 - Queste cartelle sono create dal loro software disponibile solo per Windows
 - **No SDK** disponibili
 - È necessario **leggere i file dalla cartella di esportazione**
- I dati sono salvati in un file .txt e **ogni riga rappresenta un test**
 - Ogni riga contiene id campione, id test, tipo di campione, risultato concentrazione, unità concentrazione, data risultato

Campo	N. caratteri	Valori
ID paziente	≤ 16	Stringa alfanumerica (tutti i caratteri tranne #)
ID test	≤ 16	Stringa alfanumerica (tutti i caratteri)
Tipo di campione	=3	"SER": Siero "UR": Urina "WBL": Sangue totale 'SEM': Semen 'LIQ': liquidi biologici
Risultato concentrazione	≤ 10	
Unità concentrazione	≤ 10	
Data risultato	≤ 19	gg/mm/aa h:m:s

*Analyzer A25

Architettura Dati ICE



Soluzione per interfacciamento sensori con architettura dati

- BOX-IO utilizza il sistema operativo Yocto
- Gli analizzatori sono sviluppati per funzionare con software specifici disponibili solo su Windows
- Una comunicazione diretta tra gli analizzatori e BOX-IO non è possibile



Soluzione per interfacciamento sensori con architettura dati

