

Candidato presidente
CORRADO CROCETTA

Domanda 1. Quale ritenete sia e come pensate possa evolvere il rapporto fra Statistica e Data Science?

A mio avviso un *data scientist* deve avere forti competenze di:

- a) statistica e *machine learning*
- b) *database management*
- c) gestione di sistemi di calcolo distribuito e parallelo
- d) dominio oggetto di studio
- e) comunicazione (anche mediante *data visualization*)

Statistica e *data science* hanno un comune obiettivo, ovvero l'estrazione della conoscenza dai dati ma usano approcci diversi, sia per quanto riguarda il metodo di costruzione dei modelli, le dimensioni dei database, i tipi di problemi studiati, il background culturale degli esperti ed il linguaggio utilizzato (per ulteriori approfondimenti consiglio <https://ww2.amstat.org/misc/DataScienceStatement.pdf>).

Una peculiarità del *data science* è la interdisciplinarietà, infatti, è possibile individuare 3 comunità professionali (informatici, statistici ed ingegneri) che stanno fornendo un importante contributo allo sviluppo di questo settore.

Certamente il *data science* si interseca con numerose altre discipline e aree di ricerca ma è evidente che l'apporto maggiore proviene dalla collaborazione tra queste tre comunità professionali.

La statistica e il machine learning svolgono un ruolo centrale nel *data science*. Le tecniche inferenziali e la distinzione fra causalità e correlazione permettono di estrarre conoscenza, ottenere risposte migliori e di effettuare previsioni e stime quantificando il livello di incertezza.

In sintesi, a mio avviso, gli statistici devono avere un approccio collaborativo con le comunità professionali sopra indicate, coinvolgendole in un processo di reciproco apprendimento, tenendo conto che, nei prossimi anni, ci saranno delle evoluzioni molto interessanti e rapide legate ad una forte domanda di business intelligence, digital transformation, *knowledge management*, *deep learning*, *intelligenza artificiale*, *informative data*...

Domanda 2. Quale strategia pensate possa essere utile per creare, a livello nazionale e internazionale, una solida connessione fra le comunità scientifiche che si occupano, da diverse angolazioni, di data science anche al fine di coordinare gruppi di ricerca?

A mio avviso è necessario operare su due fronti:

- a) creare delle forme di interlocuzione e di collaborazione stabili con i gruppi di informatici e di ingegneri che si occupano di questi temi. È importante invitarli nei nostri convegni ed organizzare delle sessioni nelle loro conferenze ma, soprattutto, organizzare degli eventi in collaborazione, come ad esempio il convegno DS&SR 2020 (Data Science & Social Research 2020) che il CIRPAS sta organizzando a Bari, a dicembre 2020, seguendo l'esempio delle 2 precedenti edizioni di Napoli e Milano.
- b) promuovere la pubblicazione di special issue (di riviste o di volumi) e la presentazione di progetti di ricerca su temi specifici (di natura medica, ambientale, ...) favorendo il coinvolgimento di professionalità diverse al fine di favorire le sinergie derivanti da un approccio multidisciplinare.

Domanda 3. In che modo la SIS, i suoi gruppi, e in special modo il gruppo SDS, può porsi come interlocutore privilegiato in tutte quelle nuove iniziative per la formazione dei data scientist (lauree, lauree magistrali, masters) soprattutto quando non sono direttamente organizzate dagli statistici stessi?

Come indicato nel mio programma intendo potenziare il ruolo dei gruppi SIS creando un organismo di coordinamento e prevedendo delle audizioni annuali con ciascun coordinatore di gruppo. Per incentivare la partecipazione ai convegni vorrei prevedere delle attività satellite di formazione e di terza missione organizzate dai vari gruppi. Tali attività potranno essere svolte, anche in collaborazione con le altre società scientifiche, italiane e straniere, attive nell'ambito del *data science*.

I corsi ECAS organizzati (anche online) in collaborazione con FENStatS potranno essere una ulteriore opportunità di promozione di SIS SDS nei 26 paesi europei rappresentati.

Nell'ambito delle attività di coordinamento delle attività didattiche dei dottorati di ricerca che la SIS intende attuare, il gruppo SIS SDS può svolgere un ruolo prezioso favorendo la condivisione delle informazioni relative ai vari corsi/seminari/workshop organizzati dai suoi aderenti, favorendo la partecipazione di dottorandi e di giovani ricercatori.

Nel pieno rispetto dell'autonomia organizzativa di SIS SDS, la SIS avrà un ruolo proattivo per organizzare webinar, Knowledge café e bar camp e scuole estive, per favorire il confronto su temi inerenti il DS. La SIS metterà a disposizione una bacheca multicanale per promuovere le attività organizzate da SIS SDS e per gestire archivi open per condividere i materiali didattici.

Sarebbe auspicabile che i coordinatori o i colleghi attivi nei tanti corsi di laurea e master in *data science*, creassero un gruppo di lavoro all'interno di SIS SDS per favorire lo scambio di informazioni ed esperienze e per avviare delle collaborazioni con i colleghi di altre aree disciplinari responsabili di corsi di studio e di master in *data science*.