

NON ESSERATA

CT H [signature] QB

Prova teorico-pratica C

Parte 1. Teorica

La/Il candidata/o sempre all'interno di questo Notebook, risponda alle seguenti quattro domande relative ai quattro argomenti di cui all'Art.7 comma 1 del Bando:

- 1) Tecniche di analisi dati e, in particolare, elaborazione di immagini
- 2) Progettazione e sviluppo di software scientifico in Python.
- 3) Gestione di codici in ambienti condivisi (es. Git, GitHub, GitLab). ¶
- 4) Revisione e definizione dei requisiti software in contesti collaborativi internazionali.

Theoretical-Practical Examination C

Part 1. Theoretical Section

The candidate shall answer the following four questions directly within this Notebook, concerning the four topics listed in Article 7, Paragraph 1 of the Call for Applications:

- 1) Data analysis techniques, with particular emphasis on image processing.
- 2) Design and development of scientific software in Python.
- 3) Management of code in shared environments (e.g., Git, GitHub, GitLab).
- 4) Review and definition of software requirements in international collaborative contexts.

Domanda 1) Trattamento di segnali con fondo variabile

Si consideri un'immagine, pixel vs flusso, affetta da rumore poissoniano e da un fondo continuo lentamente variabile. Descrivere un metodo per stimare il fondo in modo robusto, discutendo diversi metodi (per esempio fit polinomiale locale, spline, o altro). Evidenziare vantaggi, svantaggi dei diversi metodi e sensibilità al variare dei parametri scelti.

Question 1) Signal Processing with a Variable Background

Consider an image, represented as pixels versus flux, affected by Poisson noise and by a slowly varying continuous background. Describe a method to estimate the background in a robust way, discussing different approaches (for example, local polynomial fitting, splines, or

other methods).

Highlight the advantages and disadvantages of the different approaches, as well as their sensitivity to variations in the chosen parameters

Risposta 1)

Domanda 2

Il candidato, nella realizzazione di una pipeline Python, discuta:

- organizzazione a moduli
- uso di classi,
- gestione dei percorsi dei file e dei metadati
- logging e gestione degli errori,
- strategie per testare singoli moduli con dati simulati

Question 2)

In the development of a Python pipeline, discuss:

- modular organization;
- the use of classes;
- management of file paths and metadata;
- logging and error handling;
- strategies for testing individual modules using simulated data.

Risposta 2)

Domanda 3) Collaborazione con controllo versione e integrazione continua

In un team di più sviluppatori che lavora su un unico codice per la calibrazione di dati da telescopio, si adotta Git. Descrivere:

- una strategia di branching (es. Git Flow o GitHub Flow) adatta a sviluppi paralleli,
- come si gestiscono hotfix su codice in produzione senza interrompere nuovi sviluppi,
- come si automatizzano i test pre-merge tramite CI (es. GitHub Actions),
- la procedura di code review e merge per minimizzare conflitti.

Question 3) Collaboration with version control and continuous integration

In a team of several developers working on a single codebase for the calibration of telescope data, Git is adopted. Describe:

- a branching strategy (e.g. Git Flow or GitHub Flow) suitable for parallel development;
- how hotfixes on production code are managed without interrupting ongoing development;

- how pre-merge tests are automated through CI (e.g. GitHub Actions);
- the code review and merge procedure used to minimize conflicts.

Risposta 3)

Domanda 4) Evoluzione dei requisiti in un progetto scientifico a lungo termine

Un progetto di analisi dati per una missione spaziale ha requisiti che cambiano dopo la fase di commissioning. Descrivere come:

- raccogliere e documentare i requisiti iniziali e le modifiche (es. usando markdown + versionamento),
- tracciare la relazione tra requisiti e implementazione (matrice di tracciabilità),
- gestire le modifiche mantenendo coerenza tra specifiche, codice e test,
- strumenti suggeriti: issue tracker (GitHub Issues, Jira), versionamento semantico, test di regressione.

Question 4) Requirements evolution in a long-term scientific project

A data analysis project for a space mission has requirements that evolve after the commissioning phase. Describe how to:

- collect and document the initial requirements and subsequent changes (e.g. using Markdown documents under version control);
- track the relationship between requirements and their implementation (requirements traceability matrix);
- manage changes while maintaining consistency between specifications, source code, and tests;
- use appropriate tools such as issue trackers (e.g. GitHub Issues, Jira), semantic versioning, and regression testing.

Risposta 4)

Parte 2. Pratica

La/Il candidata/o sviluppi una routine in linguaggio Python in ambiente Jupyter Notebook, per l'analisi di un'immagine astronomica fornita dalla Commissione in formato FITS, denominata: image.fits

La routine dovrà essere sviluppata utilizzando esclusivamente librerie Python standard per l'analisi scientifica dei dati, quali numpy, matplotlib (dopo avere importato matplotlib si utilizzi il comando "%matplotlib inline" per la corretta visualizzazione delle immagini), astropy e photutils. Durante lo svolgimento della prova è consentito l'utilizzo delle funzionalità di documentazione integrate nell'ambiente di programmazione Python (docstring), accessibili localmente e senza necessità di connessione alla rete. In particolare, i candidati potranno consultare la documentazione delle funzioni e delle librerie disponibili mediante i comandi standard dell'ambiente Jupyter Notebook o del linguaggio Python (per esempio help(plt.imshow))

La/Il candidata/o dovrà:

La routine dovrà:

1. Leggere l'immagine `image.fits` e visualizzare le informazioni principali (dimensioni, tipo di dati, statistiche di base).
2. Produrre una visualizzazione con scala di contrasto ottimale (specificando la scelta dei percentili o della funzione di scaling).
3. Individuare una sorgente puntiforme (es. una stella) e definire:
 - una regione circolare per la sorgente (raggio da motivare),
 - una regione anulare per la stima del fondo cielo, evitando la sorgente stessa;
4. Stimare il fondo cielo nella regione definita al punto 3 con due metodi diversi a scelta tra:
 - mediana dei pixel,
 - sigma-clipping (descrivere parametri),
 - media. Confrontare brevemente i due risultati.
5. Calcolare il flusso lordo della sorgente (somma dei valori nella regione sorgente).
6. Sottrarre il contributo del fondo, stimare il flusso netto e il suo errore propagato.
7. Produrre una figura con:
 - l'immagine originale,
 - regione anulare attorno alla sorgente,
 - legenda.
8. Salvare in un file di testo (`.dat` o `.csv`) i seguenti risultati:
 - coordinate della sorgente,
 - flusso lordo, flusso di fondo, flusso netto,
 - errore sul flusso netto,
 - metodo usato per il fondo.
9. Commentare brevemente:
 - assunzioni fatte (es. fondo costante localmente, assenza di stelle vicine),
 - limiti del metodo (es. sensibilità al raggio, errori non considerati),
 - possibili miglioramenti.

Part 2. Practical Exercise

The candidate shall develop a Python routine in a Jupyter Notebook environment for the analysis of an astronomical image provided by the Examination Committee in FITS format, named: `image.fits`.

The routine shall be developed using only standard Python libraries for scientific data analysis, such as `numpy`, `matplotlib` (after importing `matplotlib`, use the command `%matplotlib inline` for the correct visualization of images), `astropy`, and `photutils`.

During the examination, candidates are allowed to use the documentation functionalities integrated into the Python programming environment (docstrings), accessible locally and without requiring a network connection.

In particular, candidates may consult the documentation of the available functions and libraries using the standard commands of the Jupyter Notebook environment or of the Python language, for example:

```
help(plt.imshow)
```

The candidate shall:

The routine shall:

1. Read the image `image.fits` and display the main information (dimensions, data type, basic statistics).
2. Produce a visualization with an optimal contrast scale, specifying the choice of percentiles or scaling function.
3. Identify a point source (e.g. a star) and define:
 - a circular region for the source, with the radius to be justified;
 - an annular region for the sky background estimation, avoiding the source itself.
4. Estimate the sky background in the region defined in point 3 using two different methods chosen among:
 - median of the pixels;
 - sigma-clipping, describing the adopted parameters;
 - mean.

Briefly compare the two results.

5. Compute the gross flux of the source, defined as the sum of the values within the source region.
6. Subtract the background contribution, estimate the net flux and its propagated uncertainty.
7. Produce a figure showing:
 - the original image;
 - the annular region around the source;
 - a legend.
8. Save the following results in a text file (`.dat` or `.csv`):
 - source coordinates;
 - gross flux, background flux, net flux;
 - uncertainty on the net flux;
 - method used for the background estimation.
9. Briefly comment on:
 - the assumptions made, e.g. locally constant background, absence of nearby stars;
 - the limitations of the method, e.g. sensitivity to the adopted radius, neglected sources of error;
 - possible improvements.

In []: 1