```
#!/usr/bin/env -S uv run --script

# /// script
# requires-python = ">=3.12"
# dependencies = [
#     "click",
#     "pandas",
# ]
# ///

from pathlib import Path

import click
import pandas as pd

"""
Basic Usage:
"python duplicate_removal.py -i path/to/input8/file -o path/to/output/file"

You need to have either python or uv installed in your system
You can instal uv via:
https://docs.astral.sh/uv/getting-started/installation/
"""


@click.command()
@click.option(
    "-i",
    "--input_path",
    default="input.csv",
    require=False,
    type=click.Path,
    help="Input path to file with list of papers",
)
@click.option(
    "-o",
    "--output_path",
    default="output.csv",
    required=False,
    type=click.Path,
    help="Output path with list of papers without duplicate",
)
def main(input_path: str | Path, output_path: str | Path) -> None:
    DUPLICATE_COLUMNS: list[str] = [""]  # Insert column names

    file_extension = Path(input_path).suffix

    match file_extension:
        case ".csv":
            df = pd.read_csv(input_path)
        case ".xls" | ".xlsx":
            df = pd.read_excel(input_path)
        case ".tsv":
            df = pd.read_csv(input_path, sep="\t")
```

```
        case _:  # Everything else
            raise ValueError(f"Unsopported: {file_extension}")

    df_cleaned = df.drop_duplicates(subset=DUPLICATE_COLUMNS)
    df_cleaned.to_csv(output_path, index=False)


if __name__ == "__main__":
    main()
```