Gökhan Çiflikli, Ph.D.

www.gokhan.io
GitHub | LinkedIn

Email : g.ciflikli@lse.ac.uk
Mobile : +44 7477 104110
Twitter : @gokhan_ciflikli

EDUCATION

- **London School of Economics** London, UK
  Ph.D. in International Relations *Sep. 2014 – Sep. 2018*

- **Uppsala University** Uppsala, Sweden
  Master of Social Science in Peace and Conflict Studies *Aug. 2010 – June. 2012*

EXPERIENCE

- **UCDP** (Remote)
  Data Scientist [R, Python] *Jan 2019 - Present*
    - **Natural Language Processing**: I employ industrial solutions (SpaCy, quanteda) to tokenise unstructured, multilingual text mined from aggregated news sources over a span of 30 years.
    - **Quantitative Text Analysis**: I create actionable insights from tokenised text (~48 million) using named entity recognition, POS tagging, relationship extraction, and topic modelling.
    - **Supervised Learning**: I train ensemble models of binary- and multi-classification algorithms (Naive Bayes, Random Forest, Support Vector Machines, XGBoost) to automate labelling tasks.
    - **Active Learning**: I increase the predictive accuracy of the above models (~10%) by having human annotators validate the predictions at regular intervals using the Prodigy annotation tool.

- **London School of Economics** London, UK
  Software Developer [R] *Sep 2018 - Present*
    - **Text Analysis**: I am a member of the quanteda team, a popular R package (256K downloads; 16K/month) for text analysis alongside its sister packages spacyr, readtext, and stopwords.
    - **Software**: I work on a commercial, scalable (up to 1,000 users) graphical user interface (GUI) cloud application for the quanteda package using R-Shiny, AWS, and Docker technologies.
    - **Team Management**: I manage contracted R & JavaScript programmers (3-5) using Agile/Scrum methodology via weekly video conferences and daily GitHub issue/milestone management.

- **Alan Turing Institute** London, UK
  Software Developer [R, Python] *Dec 2017*
    - **Predictive Modelling**: I built an interactive app for the UK Cabinet office for forecasting radical terrorism in the UK based on 300 profiles of convicted terrorists.

- **University College London** London, UK
  Data Scientist [R] *Sep 2015 - 2017*
    - **R package**: I designed an intuitive Random Forest fitter (RebelCast) for time-series-cross-section data to forecast escalation and severity of armed conflict which improves upon baseline models by 3%-26%.

INTERACTIVE WEB APPS [R SHINY]

- **Adavis**: I visualise political polarisation in the US Congress over the last 70 years based on voting patterns consisting of 36K roll-call votes. [App] [Code]
- **Gender Bias**: I break down the gender bias in a leading UK higher education institution reading list comprised of 13K publications finding a 80/20 divide favouring male authors. [App] [Code]

EVENTS

1st Computational Social Science (CSS) Hackathon, 17-19 April 2018 [Website]

- **Fund Raising:** As the founder of the CSS Society, I raised 10,000 GDP to fund a two-day hackathon in London, sponsored by the LSE, GitHub, Cambridge Spark, and R-Consortium.

- **Impact:** I organised workshops and talks by practitioners from Google, Alan Turing Institute, R-Ladies, Deliveroo, XL Catlin, Oxera, and Zinc to an international audience of fifty hackers.

AWARDS

- **University of California, Berkeley Bursary**: Ph.D. Mobility Program. 2016. *2,300 GBP.*

- **Michael Leifer Ph.D. Scholarship**: LSE. 2014-2017. *100,000 GBP.*

- **Facilitating Data Visualisation and Extraction: A New Interface for the UCDP**: *Collaborator.* RJ. 2014. *2.4 million SEK (~211,000 GBP).*

SKILLS

- **Programming**: R (developer), Python (advanced), SQL/Postgres (advanced), Spark

- **Statistics**: Stata, SPSS (instructor)

- **Languages**: English, Turkish   **Front-end**: HTML, CSS, LaTeX

- **Presentation**: RMarkdown, Reveal.js   **Interactive**: Shiny

Last updated on May 3, 2019