# Đề tài 1: Thu thập và phân tích dữ liệu Twitter

Với đề tài này, nhóm sẽ được thưởng thêm 1 điểm.

Yêu cầu 1: Tìm danh sách KOL trong blockchain. Gợi ý: Dùng tính năng search của Twitter, chọn 1 số hashtag/từ khóa phù hợp trong lĩnh vực blockchain, lọc ra danh sách người dùng có lượng follow/tương tác cao. Lưu ý: cần tự động hóa quá trình này, sử dụng các công cụ như Selenium.

Yêu cầu 2: Từ mỗi KOL, thu thập các tweet và một số lớn các followers của họ. Dùng thuật toán Pagerank, xây dựng đồ thị và tính điểm rank cho các node trong đồ thị, từ đó xếp hạng các KOL. Đồ thị phải có số lượng cạnh và đỉnh đủ lớn. Gợi ý: đồ thị có thể có các node như: User-Đại diện cho các KOL và user thông thường, Tweet-Đại diện cho các tweet của các KOL đó và các tweet liên quan tới crypto từ những tài khoản khác. Các cạnh kết nối: cạnh từ user này đến user khác thể hiện quan hệ theo dõi, comment, retweet, cạnh từ user đến tweet thể hiện việc đăng tweet, cạnh từ tweet đến user thể hiện việc retweet, …

# Đề tài 2: Game chơi bài

Cần xây dựng ít nhất 2 loại game với bộ bài 52 lá. Các loại game có thể là tiến lên miền Nam, miền Bắc, phỏm, 3 cây, … Có thể chọn số lượng người chơi (2, 3, hoặc 4). Có thể chơi với máy, hoặc chơi với người chơi khác (tất cả các người chơi đều chơi trên cùng 1 thiết bị).

Game có 2 chế độ hoạt động: basic và đồ họa. Cả 2 chế độ đều dùng lập trình Java GUI. Nhưng ở chế độ basic, các quân bài sẽ không có giao diện hình ảnh sinh động, mà dùng dạng text để biểu diễn.

Game phải được thiết kế tối ưu. Trước tiên phải có tính tái sử dụng: có thể dễ dàng phát triển thêm các loại hình game mới với bộ bài 52 lá. 2 chế độ basic và đồ họa cần tránh bị lặp code. Có thể dễ dàng đổi sang chế độ đồ họa mới (không chỉ là hình ảnh quân bài, mà còn nhiều yếu tố khác trên giao diện). Khi chơi với máy, không cần máy chơi thông minh. Nhưng phải thiết kế làm sao để dễ dàng thay thế các thuật toán chơi khác nhau của máy vào trong từng trò chơi.

# Cách thức nộp bài tập lớn

Các yêu cầu:

* Tất cả cần nộp cùng 1 thời điểm: upload lên thư mục GG Drive + bản cứng báo cáo. Khi bảo vệ, các nhóm download xuống
* Sản phẩm cần nộp (lưu ý, nộp các file riêng lẻ, không nộp file zip
  + Báo cáo
  + Slide trình bày
  + Video demo
  + Mã nguồn chương trình
  + Các file tài nguyên (Cần chạy chương trình ở nhà trước, để lấy được đủ dữ liệu)
  + Nộp báo cáo bản cứng
* Báo cáo:
  + Phân công công việc các thành viên trong nhóm
  + % đóng góp (con số lượng hóa cụ thể) của từng thành viên
  + Chi tiết nhật ký làm việc nhóm (sẽ được đánh giá 1đ trong điểm giữa kỳ)
  + Bảng thống kê về dữ liệu thu thập được. Ví dụ
    - Bao nhiêu loại nguồn dữ liệu, là những loại nào (news, tweets, blogs, …), mỗi loại có bao nhiêu nguồn dữ liệu, là những nguồn nào
    - Số lượng tin tức thu thập được cho mỗi nguồn
    - Thống kê về số lượng thuộc tính thu thập được cho các bài viết
  + Vẽ biểu đồ UML
    - Biểu đồ phụ thuộc gói
    - Biểu đồ lớp
  + Giải thích thiết kế
    - Các gói (package)
    - Các lớp
  + Giải thích các kỹ thuật lập trình hướng đối tượng đã áp dụng
    - Nhóm đã áp dụng kỹ thuật gì, ở đâu, lợi ích đem lại là gì
  + Liệt kê công nghệ sử dụng, thuật toán hay (nếu có)
  + Hướng dẫn sử dụng ngắn gọn + 1 số ảnh quan trọng demo chương trình