**Introduction**

Copper and copper alloys are some of the most versatile engineering materials available.(6) Pure copper has the best electrical and thermal conductivity of any commercial metal and forms alloys more freely than other metals which makes it suitable for a wide range of applications(5). The ability to freely form alloys opens a potential to get excellent physical properties such as high tensile strength, conductivity (thermal and electrical), ductility, corrosion resistance and more. These properties can be further enhanced with variations in composition and manufacturing processes.(6)

There exist more than 400 copper alloys already, each with a unique combination of properties(5). The development of new copper alloys, however, relies largely on a combination of expert judgment, trial and error and intuition, which makes the process slow and expensive(2). This indicates a requirement of system that can quickly and reliably recommend compositions and processing conditions for targeted property values.

The main objective of this study is to propose an alloy design system (referred as Smart Alloy Generation System (SAGS) from here) that utilizes machine learning to obtain multi-element copper alloy compositions and processing conditions for a desired value of either ***Tensile Strength*** or ***Thermal Conductivity*** (referred as relevant properties from here). The development of the SAGS system can be broadly divided into broad processes –

**1. Predictive Modelling**

Two distinct Random Forest models, one for tensile strength and one for thermal conductivity predictions, are built to learn the relationship between tensile strength, thermal conductivity the compositions of copper alloys and their processing conditions using samples obtained from multiple sources including public databases and educational packages.

**2. Composition Generation**

The models obtained in the previous step are used to generate candidate alloy compositions (and processing conditions in the case of tensile strength) for a user-defined value of either of the relevant properties. This is done by making property predictions on synthetically generated composition sample space using combinations of the compositions of existing alloys.

The proposed system demonstrates the ability to generate suitable compositions for a given value of the relevant properties (discussed in Chapter 5) and shows potential to alter the conventional practices of alloy discovery.

**Research Question**

**Why is it important?**

**Assumptions**

**Literature Review**

 Reddy et al.23 established an inference model from compositions and heat treatment conditions to mechanical properties of the low alloy steel by combining the back-propagation (BP) NN and genetic algorithm (GA). Their model successfully learns the influence of compositions and heat treatment conditions on the performance of the steel. Ozerdem et al.24 built a multi-layer BP NN model to predict the yield strength, UTS and elongation of the Cu–Sn–Pb–Zn–Ni alloy. These NN models with inputs of compositions and processing conditions can estimate the properties of alloys. Such forward models from composition to property is helpful to screen or down-select the potential good candidates. However, more attractive thing is an inverse design model that recommends compositions from a targeted property, i.e., a property to composition predictive model.

**Proposed solution**

**Summary of Findings**

1. Xue, D., Balachandran, P. V., Hogden, J., Theiler, J., Xue, D., & Lookman, T. (2016). Accelerated search for materials with targeted properties by adaptive design. *Nature communications*, *7*(1), 1-9.
2. Hu, X., Wang, J., Wang, Y., Li, J., Wang, Z., Dang, Y., & Gu, Y. (2018). Two-way design of alloys for advanced ultra supercritical plants based on machine learning. *Computational Materials Science*, *155*, 331-339.
3. Kim, K., Kang, S., Yoo, J., Kwon, Y., Nam, Y., Lee, D., ... & Son, W. J. (2018). Deep-learning-based inverse design model for intelligent discovery of organic molecules. *npj Computational Materials*, *4*(1), 1-7.
4. Raccuglia, P., Elbert, K. C., Adler, P. D., Falk, C., Wenny, M. B., Mollo, A., ... & Norquist, A. J. (2016). Machine-learning-assisted materials discovery using failed experiments. *Nature*, *533*(7601), 73-76.
5. Rovinelli, A., Sangid, M. D., Proudhon, H., & Ludwig, W. (2018). Using machine learning and a data-driven approach to identify the small fatigue crack driving force in polycrystalline materials. *npj Computational Materials*, *4*(1), 1-10.